

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Vid Bevčar

Optimizacija parametrov umerjanja cevovodnega sistema

MAGISTRSKO DELO
ŠTUDIJSKI PROGRAM DRUGE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: prof. dr. Igor Kononenko
SOMENTOR: prof. dr. Franc Steinman

Ljubljana, 2016

Rezultati magistrskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov magistrskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Za pomoč pri izdelavi naloge in usmerjanje se zahvaljujem mentorju prof. dr. Igorju Kononenku in somentorju prof. dr. Francu Steinmanu. Prav tako se zahvaljujem dr. Danielu Kozelju za vabilo v Laboratorij za mehaniko tekočin, idejo za nalogo ter seveda obširno pomoč pri implementaciji in pisanju.

Zahvaljujem se tudi družini in puncu Heleni, ki so mi stali ob strani in me podpirali med študijem.

Posvetilo

*Magistrsko nalogo posvečam pokojnemu očetu
Igorju.*

Kazalo

Povzetek

Abstract

1	Uvod	1
1.1	Motivacija in cilji	2
1.2	Struktura naloge	3
2	Hidravlični modeli	5
2.1	Uvod	5
2.2	Gradniki vodovodnega sistema	6
2.3	Osnove hidravlike	8
2.4	Izračun hidravličnega modela	13
2.5	Knjižnica EPANET	23
3	Genetski algoritmi	27
3.1	Metodologija	27
3.2	Pareto optimalne rešitve	28
3.3	Delovanje NSGA-2	34
3.4	Genetski operatorji	36
4	Umerjanje parametrov vodovodnih sistemov	41
4.1	Uvod	41
4.2	Umerjanje z genetskimi algoritmi	45
4.3	Analiza kalibracije	56

KAZALO

5	Umerjanje na primerih	71
5.1	Priprava meritev	71
5.2	Uporabljeni modeli	72
5.3	Primer: Anytown AM	75
5.4	Primer: Aques de Coimbra	81
5.5	Primer: Aques de Coimbra z odstranjevanjem grup	88
6	Implementacija kot GIS aplikacija	93
6.1	Uvod	93
6.2	Vrste podatkovnih tipov	93
6.3	Vzpostavitev hidravličnega modela	95
6.4	QGIS	95
6.5	Vtičnik CALIB	96
7	Zaključek	99
7.1	Nadaljnje delo	100

Seznam uporabljenih kratic

kratica	angleško	slovensko
GA	genetic algorithm	genetski algoritem
VS	water distribution system	vodovodni sistem
GIS	Geographical information system	Geografski informacijski sistem

Povzetek

Naslov: Optimizacija parametrov umerjanja cevovodnega sistema

Parametri vodovodnega sistema potrebujejo, podobno kot vsi matematični modeli, kalibracijo, preden lahko iz njih pridobimo uporabne rezultate. V tej nalogi je opisana kalibracija na osnovi večkriterijske optimizacije. Parametre najprej grupiramo na različne načine, da povečamo umerljivost modela. Optimizacijo opravimo s pomočjo genetskih algoritmov, ki so v nalogi podrobno opisani. Ker je negotovost parametrov pomemben del umerjanja, smo v ta namen uporabili standardno metodo FOSM in metodo na podlagi statistične analize populacije. Poleg tega smo uporabili številna dodatna orodja za verifikacijo rezultatov in ujemanja modela. Kalibracijski algoritem kaže dobre rezultate, tako v izmišljenem kot tudi v realnem primeru kalibracije vodovodnega sistema. Metodi določanja negotovosti pokazeta nekoliko slabše, a še vedno zelo uporabne rezultate.

Ključne besede: vodovodni sistem, umerjanje, večkriterijska optimizacija

Abstract

Title: Optimisation of water distribution system calibration parameters

Water distribution system (WDS) model parameters, as all mathematical models, need calibration, before practical results can be retrieved. In this thesis a multiobjective optimization approach to calibration is described. Firstly different parameter groupings are used to improve calibration stability. A genetic algorithm based optimization algorithm is used and in depth described. Since parameter uncertainty is an important part of calibration process, a standard FOSM method and a GA population statistics based method are used in the process. In addition, the analysis tools for calibration and model fitting validation are described. Calibration algorithm shows promising results when applied to hypothetical and real life calibration scenarios. Uncertainty analysis shows sub optimal, but still very useful results.

Keywords: water distribution system, calibration, multiobjective optimization

Poglavje 1

Uvod

Dostop do zanesljive, varne in cenovno ugodne pitne vode je ključnega pomena za vsako skupnost in je pogoj za učinkovit nadaljnji razvoj in višanje kvalitete življenja. Zagotavljanje omenjenih pogojev ob hkratnem povečevanju naselij, pa tudi načrtovanje novih, narekuje uporabo orodij za analizo delovanja vodovodnih sistemov [1].

Hidravlični modeli vodovodnih sistemov se uporabljajo tako med projektiranjem kot med obratovanjem sistema. Omogočajo nam virtualno simulacijo kakovosti vode in njenih hidravličnih lastnosti, pomagajo pri projektiranju in obratovanju ter načrtovanju obnov.

Kot pri vseh matematičnih modelih potrebujejo tudi parametri hidravličnih modelov kalibracijo, preden lahko iz njih dobimo praktično uporabne informacije [2]. Kljub temu pride, zaradi omejitev, pri zajemu podatkov iz vodovodnega sistema do nezanemarljivih odstopanj med modelom in situacijo v naravi [3].

Glavna tema te naloge je umerjanje hidravličnih modelov vodovodnih sistemov in razvoj orodij za analizo rezultatov umerjanja. Vodovodni sistem predstavimo kot neusmerjen graf, kjer so povezave cevi in črpalke, vozlišča pa spoji in rezervoarji. Na povezavah ponavadi merimo pretoke, na vozliščih pa tlake. Z leti uporabe se elementom spremenijo nekatere lastnosti, kot na primer cevem hrapavost, ki povečuje hidravlične izgube. Te lastnosti so ponavadi težko izmerljive. Tekom umerjanja poskušamo posameznim elementom sistema določiti te lastnosti (parametre), da se rezultati meritev na terenu dobro skladajo s hidravličnim modelom [4].

1.1 Motivacija in cilji

Samoumevno izgleda, da je dostop do kvalitetne pitne vode ena od osnovnih potreb vsakršne moderne civilizacije. Precej manj samoumevna pa je raba hidravličnih modelov vodovodnih sistemov. Posebej zato, ker se v sodobnem času ne zavedamo, kakšna infrastruktura je potrebna za dostavo te življenjske dobrine.

Danes se hidravlični modeli uporabljajo v različne namene. Pri načrtovanju novih in obnovah obstoječih omrežij si inženirji z njimi pomagajo, da zadostijo vrsti kriterijev. Omrežje mora zagotavljati zadostne količine in tlake vode v primeru požara na hidrantih. Povsem druge vrste študije se na primer osredotočajo na dolgoročni razvoj naselja in posledično dimenzioniranje vodovodnega sistema. Hidravlične modele uporabljamo tudi kot osnovo modelov kvalitete vode, ki naprimer simulirajo startost vode in prenos snovi od vodnega vira po omrežju. Z njihovo pomočjo se tudi načrtuje obratovanje črpalk in na ta način zmanjšuje obratovalne stroške. Pomagajo pri odločitvah, kjer preverjamo, kako bi sanacija tega ali onega odseka vplivala na skupno obratovanje sistema.

Preden lahko hidravlični model uporabimo za katero od navedenih nalog, je potrebno parametre modela najprej umeriti, torej najti takšne vrednosti, da model posnema situacijo v naravi. V nasprotnem primeru je model bolj ali manj neuporaben, saj vrača napačne informacije.

Cilj te naloge je razvoj orodij za umerjanje parametrov vodovodnega omrežja in bralcu iz računalniške stroke približati področje mehanike tekočin v vodovodnih sistemih. Kot je razvidno v nadaljevanju, je potrebna za kvalitetno umerjen hidravlični model vrsta algoritmov. Parametre najprej razvrstimo v grupe in tako povečamo umerljivost. Sledi postopek umerjanja, na koncu pa še verifikacija rezultatov. V nalogi je celoten postopek opisan, implementiran in demonstriran na primerih. Vsa navedena orodja so implementirana znotraj geografske informacijskega sistema in tako omogočajo lažje testiranje, verifikacijo in vizualizacijo rezultatov uporabljenih algoritmov.

Naloga je namenjena računalniški stroki. Ker pa gre za interdisciplinarno področje, je posebna pozornost namenjena pregledu mehanike tekočin v vodovodnih sistemih.

1.2 Struktura naloge

V poglavju 2 postavimo temelje mehanike tekočin v tlačnih sistemih. Poglavje je namenjeno predvsem računalniški stroki, ki nima predznanja o omenjenem področju. Najprej so opisani osnovni gradniki vodovodnega sistema in temeljni hidravlični zakoni, ki so relevantni za tlačne hidravlične sisteme. Opisane so enačbe, ki zakone prenesejo v matematični svet in načini zapisa sistemov enačb, ki opišejo celoten vodovodni sistem. Pregledani so pristopi reševanja omenjenih sistemov enačb in podrobneje opisan sistem z gradientno metodo, ki je uporabljen v sklopu te naloge. Poseben razdelek je namenjen časovni analizi omenjene metode, katere ni moč zaslediti v nobeni literaturi. Na koncu je opisana knjižnjica EPANET za hidravlične simulacije.

Poglavje 3 je namenjeno povsem drugemu področju. Tokrat se spustimo v računalniške vode, natančneje v večkriterijsko optimizacijo in reševanje le-te z genetskimi algoritmi (GA). Na začetku je opisan pristop k reševanju problemov z GA. V nadaljevanju je večina pozornosti namenjena algoritmu NSGA-II, ki je uporabljen v sklopu te naloge. Opisani so njegovi ključni elementi in opravljene izpeljave časovnih zahtevnosti.

V poglavju 4 povežemo znanje o hidravličnih modelih iz poglavja 2 in znanje o večkriterijski optimizaciji iz poglavja 3 in tako dobimo algoritem za umerjanje parametrov vodovodnega sistema. S pomočjo algoritma NSGA-II poiščemo takšne parametre, da se hidravlični model čim bolj približa situaciji v naravi. Pred tem utemeljimo optimalnost takšnega pristopa s pregledom ostalih metod umerjanja. Najprej je predstavljena povezava med GA in hidravličnimi modeli. Nato je razdelek namenjen grupaciji cevi v večje, bolj umerljive skupine, na koncu pa povežemo vse skupaj v smiselno celoto. Drugi del poglavja je namenjen analizi, ki jo opravimo po umerjanju. Opisani so pristopi ugotavljanja stopnje ujemanja, ki povedo, kolikšno je ujemanje modela s situacijo v naravi. Posamezne parametre analiziramo in ugotovimo, kolikšna je njihova umerljivost, torej koliko lahko zaupamo v njihovo vrednost. V ta namen sta bili implementirani dve, med seboj zelo različni metodi. Na koncu je razdelek namenjen analizi odstopanj med meritvami in hidravličnim modelom.

Poglavje 5 je namenjeno meritvam, torej prikazu praktičnega delovanja vseh orodij na vodovodnih sistemih. Prikazani so trije primeri, najprej umerjanje

manjšega obvladljivega modela vodovodnega sistema. Sledijo dva primera umerjanja večjega, realnega sistema, v katerem je približno 29km cevovodov.

Zadnje poglavje je namenjeno lažji, bolj inženirski tematiki. Vsa omenjena orodja so bila v tej nalogi implementirana znotraj geografsko informacijskega sistema (ang. Geographical information system) na kratko GIS. V tem poglavju so opisane osnove GIS. V nadaljevanju je opisano okolje QGIS, vtičnik Oslandia za vzpostavitev hidravličnih modelov in vtičnik CALIB, znotraj katerega so vsebovani algoritmi, opisani v tej nalogi.

Poglavje 2

Hidravlični modeli

2.1 Uvod

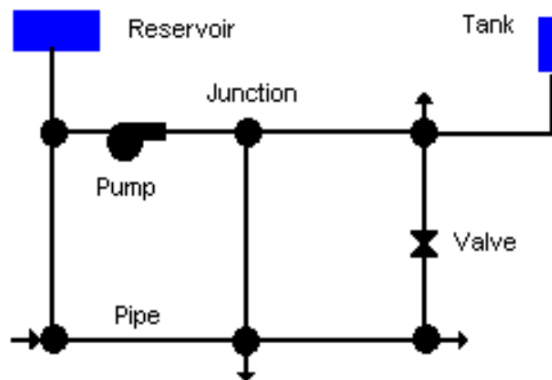
Hidravlični modeli vodovodnih sistemov so se prvič pojavili v štiridesetih letih prejšnjega stoletja, več kot dvajset let kasneje pa so luč sveta ugledali prvi računalniški modeli. Shamir in Howard sta izdelala prvega v letu 1968 ter kasneje Epp in Fowler v letu 1970.

Kasneje so bili razviti številni drugi programi, ki izkoriščajo tako nepredek v numerični matematiki kot tudi Moorovo pravilo in boljše ostalo strojno opremo. Verjetno je najbolj prepoznavno orodje EPANET, ki ga razvija Ameriška agencija za varstvo okolja (ang. United States Environmental Protection Agency [5]) na kratko EPA. Druga tovrstna orodja so še UNIWB-LOOP, WADISO, U of K KYPIPE in WATER [6].

V tem poglavju so najprej opisani osnovni gradniki vodovodnega sistema, preko katerih prav vsak od nas prejema pitno vodo. V nadaljevanju je pregled osnovnih fizikalnih zakonov, ki opisujejo gibanje tekočin po vodovodnih sistemih in predstavitev ustreznih matematičnih formulacij za njihov zapis. Našteti so načini reševanja hidravličnih simulacij. Podrobneje je gradientna metoda, ki je uporabljena znotraj te naloge. Na koncu je še opisana knjižnjica EPANET za delo s hidravličnimi modeli.

2.2 Gradniki vodovodnega sistema

vodovodni sistem si najlažje predstavljamo kot neusmerjen ciklični graf z vozlišči in povezavami med njimi. Tudi knjižnjica EPANET, katero bomo uporabljali za hidravlične izračune, uporablja predstavitev z grafi. Na sliki 2.1 je viden primer takšnega omrežja in njegova predstavitev. Vozlišča grafa predstavljajo:



Slika 2.1: Predstavitev vodovodnega sistema z neusmerjenim grafom [5].

1. Spoji

Spoji so stičišča dveh ali več povezav, ponavadi cevi, lahko pa so tudi končni elementi - porabniki. Osnovni podatki spoja so njegova nadmorska višina in poraba. V bolj naprednih modelih se poraba spreminja s časom, lahko pa je tudi negativna, kjer voda vstopa v sistem. Neznana fizikalna količina v večini spojev je tlak, katerega pridobimo tekom hidravličnega izračuna.

2. Vodni viri

Vodni viri - rezervoarji - predstavljajo neomejene vodne vire, kot so reke, jezera, izviri ali povezave z drugimi vodovodnimi sistemi. Osnovni atributi so nadmorska višina, pritisk (če je vir pod pritiskom) in količina vode. Ker je začetni element hidravličnega modela, so njegove lastnosti fiksne in neodvisne od dogajanja v modelu. Kot taki bodo služili kot referenčna točka tekom hidravličnega izračuna.

3. Vodohrani

Vodohrani so objekti, katerih glavna naloga je shranjevanje tekočine. Volumen tekočine se spreminja s časom tekom simulacije ali obratovanja. V praksi so to ponavadi cilindrične ali pravokotne zgradbe s konstantnim presekom. V vodovodnem omrežju imajo podobno lastnost kot kondenzator v električnem vezju - shranjujejo in oddajajo energijo. V vodohranu se z višanjem gladine potencialna energija zvišuje, podobno kot v kondenzatorju preko višanja naboja. Njihove osnovne lastnosti so nadmorska višina, premer, minimalna in maksimalna gladina in začetna gladina (gladina ob inicializaciji modela). Tekom simulacije lahko za vodohran izračunamo višino gladine. Minimalna in maksimalna višina gladine sta pomembni točki v algoritmu za izračun razširjene časovne analize, kot je opisano v nadaljevanju.

Poleg vozlišč graf sestavljajo tudi povezave. Literatura jih deli še na dve podvrsti, t.i. linijske elemente in virtualne linije. Oboji predstavljajo realne fizične komponente v omrežju, pri čemer prvi razpenjajo fizikalno nezanemarljive razdalje med vozlišči. Virtualne linije po drugi strani predstavljajo elemente, ki močno vplivajo na spremembe energije v vodovodnem sistemu, pri čemer so le-ti v realnosti točkovni elementi.

1. Cevi

Cevi so povezave med posameznimi vozlišči v grafu. EPANET privzema, da so vse cevi polne ves čas. Smer pretoka vode je od vozlišča z višjim pritiskom proti tistemu z nižjim. Vsaki cevi tako določimo attribute, kot so začetno, končno vozlišče, premer, dolžina in koeficient hrapavosti, ki igra pomembno vlogo v nadaljevanju. Cevi lahko določimo tudi status (odprta, zaprta) in s tem simuliramo zaprtje cevi.

Tekom hidravličnega izračuna se cevi določijo neznane fizikalne vrednosti: hitrost toka in z njo povezan pretok, izguba tlaka in Daicy - Weibachov koeficient trenja.

2. Črpalke

Črpalke v grafu predstavimo kot povezave, štejemo pa jih med virtualne linije. Njihova realna dolžina je zanemarljiva. Vodi dodajajo energijo in s

tem povečujejo pritisk, kot je podrobneje opisano v nadaljevanju. Atributi črpalke so vhodno in izhodno vozlišče ter njena moč.

3. Ventili

Ventili so povezave, ki omejujejo tlak ali pretok čez neko povezavo v omrežju. Kot attribute jim določimo začetno in končno vozlišče, premer in nastavitve ventila. Obstaja več tipov ventilov, v katere pa se ne bomo spuščali.

Poleg fizičnih komponent, ki sestavljajo neusmerjen graf, so gradniki hidravličnega modela vodovodnega sistema tudi nefizične komponente. Za natančen popis hidravličnih razmer stroka uporablja vrsto elementov, ki se ne pojavljajo na zemljevidih a opisujejo delovanje fizičnih elementov. V sklopu te naloge naštejemo le nekaj najpomembnejših.

Karakteristične krivulije predstavljajo zvezo med dvema fizikalnima količinama. Krivulija črpalke tako opisuje zvezo med črpalno višino in pretokom črpalke. Ob isti moči lahko na višjo lego prečrpamo manj tekočine kot na nižjo ob nespremenjenih ostalih pogojih [7].

Drugi tip krivulije - krivulija porabe - podaja zvezo med porabo in časom. Kot nam je intuitivno znano, je poraba večine gospodinjstev večja v jutrajnjih in popoldanskih urah, manjša pa ponoči. Z omenjeno krivulijo lahko naprimer opišemo gibanje dnevne porabe.

Pomembne nefizikalne komponente so še pravila. Slednja ob določenih izpolnjenih pogojih vplivajo na delovanje drugih elementov vodovodnega sistema. Na ta način lahko naprimer vklopimo črpalko, če se gladina vode v vodohranu spusti pod določen nivo.

2.3 Osnove hidravlike

Ključne gradnike vodovodnega sistema smo našeli, sedaj pa si pogledjmo dva fizikalna zakona, ki opisujeta dinamiko tekočin in sta ključna za nadaljnje delo.

2.3.1 Zakon o ohranitvi mase

Zakon o ohranitvi mase se nanaša na vozlišča grafa. Privzema, da je voda nestisljiva tekočina in poenostavlja, da se njena gostota ne spreminja s temperaturo

[7]. Tako lahko rečemo, da je vsota pritekajočih mas v vozlišče enaka vsoti iztekajočih. Pravilo je za vodovodna omrežja podobno kot Kirchoffovo prvo pravilo [8] v električnih vezjih. Zakonu v nadaljevanju rečemo tudi kontinuitetna enačba.

$$\sum_{vhodne} m_i = \sum_{izhodne} m_j \quad (2.1)$$

2.3.2 Zakon o ohranitvi energije

Zakon o ohranitvi energije oziroma njegova poenostavitev velja samo za sisteme s stalnim tokom. Voda v dovolj kratkem odseku enostavne cevi s po enim vhomom in iztokom nosi tri tipe energije:

1. Potencialna je energija, ki jo ima tekočina zaradi višine, na katero je postavljena. Izračunamo jo kot $E_p = m * g * h$, pri čemer je m masa, g gravitacijski pospešek in h višina, na katero je tekočina postavljena.
2. Kinetična je energija, ki jo ima tekočina zaradi svojega gibanja s hitrostjo v . Popiše jo enačba $E_k = \frac{1}{2}mv^2$, kjer je v hitrost in m masa.
3. Notranja (E_n) je energija, ki jo ima tekočina zaradi svojega položaja v cevovodu. Voda, ki leži nad obravnavanim odsekom, deluje z določeno silo, katere posledica je tlak v cevovodu. Posledica omenjenega tlaka je notranja energija.

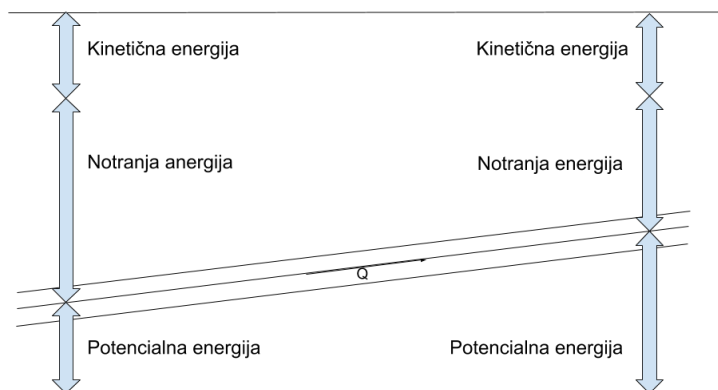
Skupna energija vode v "dovolj" kratkem odseku cevi je tako enaka vsoti vseh treh komponent:

$$E = E_p + E_k + E_n = m * g * h + \frac{1}{2}m * v^2 + E_n \quad (2.2)$$

Po prvem zakonu termodinamike vemo, da je sprememba energije v sistemu enaka dovedeni oziroma odvzeti energiji zaradi opravljenega dela. Zapišemo lahko enačbo:

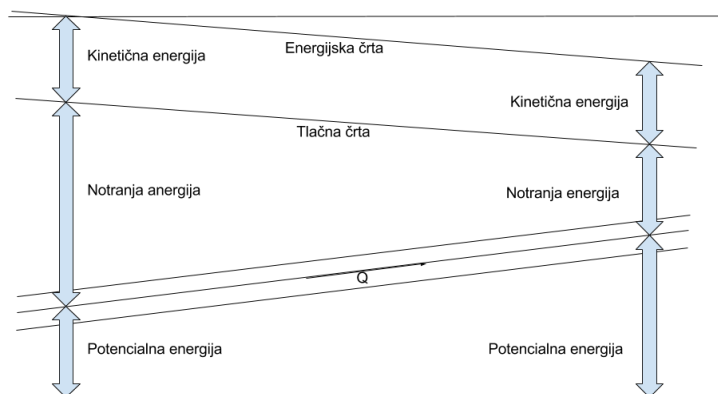
$$\Delta A = E_1 - E_2 \quad (2.3)$$

V brezizgubnih razmerah se energija znotraj sistema ohranja. Gibajoča tekočina ohranja svojo kinetično energijo, notranja (tlačna) energija pa se z dviganjem pretvori v potencialno, kot je prikazano na sliki 2.2.



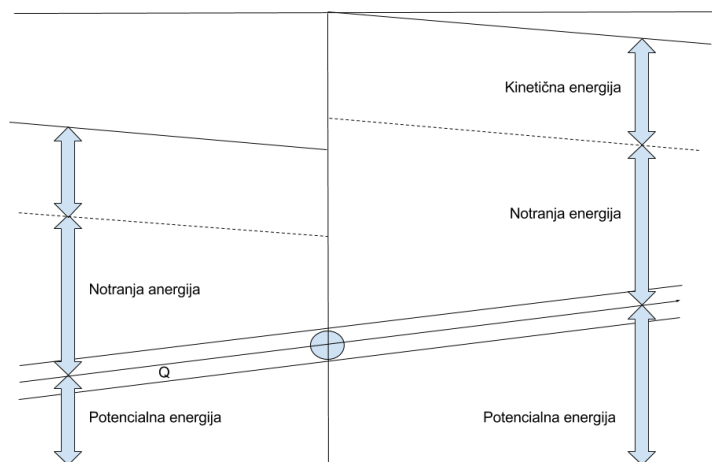
Slika 2.2: Brezizgubni sistem.

V praksi opazimo, da tekočina med gibanjem po cevi del svoje energije odda v obliki mehanskega dela. Zaradi trenja se del energije pretvori, energija na iztoku pa je manjša od energije na vtoku, kot je vidno na sliki 2.3.



Slika 2.3: Sistem z izgubo.

Energijo lahko v cevovod tudi dovajamo s pomočjo črpalke oziroma odvajamo s pomočjo redukcijskega ventila. Posledica takšnega elementa je zvišanje kinetične in tlačne energije, potencialna pa se na elementu ne spremeni (glej sliko 2.4).



Slika 2.4: Sistem z izgubo in črpalko.

Intuitivno smo spoznali zakon o ohranitvi energije. Izpeljavo energijske enačbe, ki opisuje tok v cevovodnih sistemih, smo zato kar izpustili in napisali energijsko enačbo za tok tekočine vzdolž ene osi cevovoda. Oblika enačbe meri energijo v metrih vodnega stolpca, kar je z inženirskega stališča bolj smiselno [7].

$$\frac{p_1}{\gamma} + z_1 + \frac{v_1^2}{2 * g} + h_m = \frac{p_2}{\gamma} + z_2 + \frac{v_2^2}{2 * g} + h_l \quad (2.4)$$

Kjer so posamezni členi obrazloženi v tabeli 2.1.

Člen	Pomen
$\frac{p_1}{\gamma}$	Tlačna energija
z_1	Potencialna energija
$\frac{p_1}{\gamma} + z_1$	Statična energija (tlačna črta)
$\frac{v_1^2}{2 * g}$	Kinetična energija
$\frac{p_1}{\gamma} + z_1 + \frac{v_1^2}{2 * g}$	Skupna energija (energijska črta)
h_m	Dodana ali odvzeta mehanska energija
h_l	Energijske izgube

Tabela 2.1: Členi energijske enačbe.

Opazimo, da se fizikalne enote posemaznih členov vedno okrajšajo v dolžinsko enoto - meter. Od tu naprej se vse oblike energije merijo v metrih vodnega stolpca. Sedaj lahko tudi povežemo prej razložene energijske diagrame z energijsko enačbo. Če iz enačbe izvzamemo člen, ki dodaja mehansko energijo h_m in člen energijskih izgub h_l , dobimo diagram na sliki 2.2. Če upoštevamo samo člen z izgubami, energije pa ne dovajamo, dobimo sistem z diagrama 2.3, ki opisuje gibanje tekočine po cevi. Če pa se energija na cevovodu hkrati izgublja in dodaja, dobimo sistem na diagramu 2.4, ki opisuje črpalko.

2.3.3 Energijske izgube

Omenili smo že, da se energija tekočine med premikanjem izgublja, sedaj pa si pogledjmo vzroke za ta pojav in izračun padca vodnega stolpca kot posledico. Energijske izgube delimo v dve skupini:

1. Linijske so izgube, do katerih pride med gibanjem tekočine na daljše razdalje. Med premikanjem nastaja trenje tako med delci tekočine kot med tekočino in steno cevi. Izgube, kot izgubo energije ali znižanje vodnega stolpca, opisuje Daicy-Weisbachova enačba, za znano dolžino cevi, hitrost in nestisljiv fluid izračunamo izgubo kot:

$$h_L = \lambda \frac{L * v^2}{d * 2 * g} \quad (2.5)$$

V enačbi je L dolžina cevi, d premer, λ pa brezdimenzijski Daicy-Weisbachov koeficient trenja. V enačbi je v hitrost tekočine, g pa gravitacijski pospešek. Izračun linijskih izgub je sicer zelo obširno poglavje mehanike tekočin, v podrobnosti katerega se nismo spuščali.

2. Lokalne so izgube energije ob nenadnih spremembah toka tekočine. Pojavljajo se na spojih, ventilih, kolenih in zasunih v cevovodu. Podobno, kot pri linijskih izgubah, tudi lokalne opišemo z padcem energije oziroma vodnega stolpca. Opiše jih enačba:

$$h_l = \xi_{tok} \frac{v^2}{2 * g} \quad (2.6)$$

Analitični popis tovrstnih pojavov je zaradi turbulentnega toka in kavitacije kompleksen in pogosto nemogoč. Koeficient lokalnih izgub ξ_{tok} se zato pogosto določi eksperimentalno. Odvisen je od geometrije in lastnosti tekočine [7].

2.4 Izračun hidravličnega modela

2.4.1 Formulacija enačb

Pri hidravličnem modelu vodovodnega sistema s stalim tokom si pri reševanju pomagamo z enačbo zakona o ohranitvi mas, ki se nanaša na vozlišča grafa in energijsko enačbo zakona o ohranitvi energije, ki se nanaša na povezave grafa. Množici vozlišč in povezav tako tvorita sistem nelinearnih enačb, s pomočjo katerih lahko ugotovimo neznanke modela, pretoke v ceveh in tlake v vozliščih. Na omenjen način lahko rešimo tudi časovno diskretne simulacije s spremenljivimi pretoki, t.i. razširjeno časovno simulacijo. Slednja potrebuje še dodatno informacijo o spremembah gladin v rezervoarjih zaradi spremenljivih iztokov in vtokov v sistem [9]. V tem razdelku se osredotočamo le na sisteme s stalnim tokom.

Za potrebe predstavitve enačb v sistemu nelinearnih enačb smo enačbe omenjenih zakonov še enkrat napisali v nekoliko prilagojeni obliki:

1. Zakon o ohranitvi mase

Enačbo iz razdelka 2.3.1 preoblikujemo v obliko:

$$\sum_{vhodne} Q_{vhodni} - \sum_{izhodne} Q_{izhodni} = q_i \quad (2.7)$$

Q_{vhodni} so pritekajoči pretoki, $Q_{izhodni}$ iztekajoči, razlika q_i pa je zunanja poraba, iztekajoča iz sistema. Na ta način pokrijemo tudi situacije, kjer je vozlišče list grafa in predstavlja porabnika [9].

2. Zakon o ohranitvi energije

Na poti med vozliščima A in B , ki jo sestavljajo samo cevi, lahko zakon o ohranitvi energije zapišemo kot:

$$H_A - H_B = \sum_{i \in l_{path}} h_{l,i} \quad (2.8)$$

Tako smo formulo iz razdelka 2.3.2 generalizirali za zaporedje večih cevi [9]. V formuli sta H_A in H_B tlaka v vozliščih A in B , $h_{l,i}$ pa energijske izgube na poti. Formula je uporabna za samo eno cev, zaporedje cevi in zanko, kjer je l_{path} množica cevi na poti od A do B .

Formula povzema, da je razlika tlakov v dveh vozliščih A in B , med katerima je zaporedje cevi l , enaka vsoti energijskih izgub posameznih cevi $h_{l,i}$.

3. Zanke

Zaprte zanke so zaporedja povezav (cevi), ki se začnejo in končajo v istem vozlišču (spoju). Ker gre pri začetnem in končnem vozlišču za isti spoj, je vsota energij po zaprti zanki enaka 0, zato zapišemo enačbo:

$$H_A - H_A = 0 = \sum_{i \in l_{path}} h_{l,i} \quad (2.9)$$

Psevdo zanke so poti cevi med dvema vozliščima z znanim energijskim nivojem, na primer med dvema vodnima viroma. Ker v psevdo zankah ni fizične povezave med npr. dvema vodnima viroma, se kot razlika energij uporabi kar razlika med energijskima nivojema teh dveh elementov.

2.4.2 Formulacija sistema enačb

Neznanke v hidravlični simulaciji vodovodnega sistema s stalnim tokom so pretoki v vseh ceveh Q_i in tlaki v vozliščih H_i . V sistemu z N vozlišči in M povezavami imamo $N + M$ neznak. Prej omenjene enačbe lahko formuliramo na štiri načine, ki rešijo sistem enačb. Pristope na tem mestu samo naštejemo, opišemo pa le t.i. enačbe povezav. Ta način zapisa uporablja knjižnjica EPANET, uporabljena v tej nalogi.

1. **Enačbe zank**
2. **Enačbe vozlišč in zank**
3. **Enačbe vozlišč**
4. **Enačbe povezav**

Največji sistem enačb sestavlja $N_{povezav}$ enačb cevi, ki opisujejo zakon o ohranitvi energije in $N_{vozlisce}$ enačb spojev, pri katerih se upošteva ohranitev mase. Neznanke so tlaki vozlišč H in pretoki cevi Q . Tak način zapisa hidravličnega modela je večji od prej opisanih, a omogoča učinkovitejše reševanje sistema enačb [7].

2.4.3 Metode reševanja sistema nelinearnih enačb VS

V tem razdelku le naštejemo metode reševanja sistemov nelinearnih enačb VS. Podrobneje je opisana Newton - Raphsenova metoda, iz katere izhaja gradientna metoda, uporabljena v tej nalogi.

1. **Hardy-Crossova metoda**
2. **Linearizacijska metoda**
3. **Newton - Raphsenova metoda**

Newtonova oziroma Newton - Raphsenova metoda je pogost način numeričnega iskanja ničle funkcije, kjer na vsakem koraku iščemo boljši približek x , funkcije $f(x) = 0$. Metoda začne v začetnem približku x_0 , za delovanje pa potrebujemo še funkcijo f in njeno odvod f' . Če je f odvedljiva in začetni približek x_0 dovolj dober, poiščemo boljši približek x_{i+1} po formuli:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \quad (2.10)$$

Geometrijsko predstavlja $(x_i, 0)$ presečišče med abscisno osjo in in tangento na graf funkcije f v točki $(x_0, f(x_0))$. Z iteracijo zgornje formule se zmanjšuje napaka, dokler ni dosežen zaustavitveni pogoj $x_{i+1} - x_i < \epsilon$.

Z Newton - Raphsenova metodo rešimo tudi sistem k nelinearnih zvezno odvedljivih enačb $F : R^k \mapsto R^k$. Reševanje sistema nelinearnih enačb se pretvori v reševanje sistema linearnih enačb:

$$J_F(x_n)(x_{n+1} - x_n) = -F(x_n) \quad (2.11)$$

V sistemu je $J_F(x_n)$ $k \times k$ Jakobijeva matrika, $(x_{n+1} - x_n)$ pa predstavljajo neznanke[18].

Prednost Newton - Raphsenove metode je hitra konvergenca. Reši vse zapise enačb hidravličnega modela.

Med načini reševanja in zapisa sistemov nelinearnih enačb so precejšnje razlike. Največje razlike so v hitrosti konvergence, kot je razvidno iz tabele 2.2. Meritve so povzete po [10] merijo število iteracij do doseženega konvergenčnega kriterija.

Metoda reševanja	Število iteracij
Hardy - Cross	635
Linearizacijske metode	4
Newton - Raphsen	24

Tabela 2.2: Število interacij pri različnih metodah reševanja sistema nelinearnih enačb.

2.4.4 Reševanje z gradientno metodo

Gradientna metoda je bila razvita za reševanje hidravličnih enačb iz Newton - Raphsenove metode. Uporablja jo knjižnjica EPANET, orodje za hidravlične simulacije vodovodnih sistemov.

Hidravlično simulacijo predstavimo s topologijo sistema, parametri sistema in karakteristikami gradnikov VS. Parametri so naprimer porabe po vozliščih q , fiksni vodni stolpci (pri vodnih virih) h in pravila za energijske izgube po ceveh $K(Q)$. Radi bi določili neznane pretoke v ceveh Q in tlake v vozliščih h . Enačbe smo

v nadaljevanju obravnavali v matrični obliki. Energijske enačbe se zapišejo kot sistem linearnih enačb:

$$A_{12}h + K(Q) = -A_{10}h_0 \quad (2.12)$$

Kontinuitetne enačbe pa kot sistem enačb:

$$A_{21}Q = q \quad (2.13)$$

V enačbah so posamezne komponene:

Element	Opis
$A_{12} = A_{21}^T$	Incidenčna matrika neznanih tlakov velikosti (nc,nn)
$A_{10} = A_{01}^T$	Incidenčna matrika znanih vodnih stolpcev velikosti (ne,no)
$Q^T = [Q_1, Q_2, \dots, Q_{ne}]$	Pretoki v ceveh
$q^T = [q_1, q_2, \dots, q_{nn}]$	Porabe v vozliščih
$h^T = [h_1, h_2, \dots, h_{nn}]$	Tlaki v vozliščih
$h_0^T = [h_{01}, h_{02}, \dots, h_{0nn}]$	Znani vodni stolpci v vozliščih
$K^T(Q) = [K_1, K_2, \dots, K_{ne}]$	Pravila, ki opisujejo izgube v ceveh
nn	Število vozlišč z neznanim tlakom
no	Število vozlišč z znanim tlakom
ne	Število cevi z neznanim pretokom

Incidenčne matrike A_{12} in A_{01} so sestavljene iz samih enk in ničel in posamezne vrednosti imajo naslednje pomene:

1	Voda priteka iz cevi i v vozlišče j
0	Cev i in vozlišče j nista povezani
-1	Voda izteka iz vozlišča j v cev i

Sistemi linearnih in nelinearnih enačb 2.12 in 2.13 imajo lahko več kot eno rešitev. Dokazano je, da v kolikor so $K(Q)$ monotono naraščajoče funkcije, je

rešitev omenjenega sistema enolična [10]. Todini in Pilati sta kot posledico pokazala, da lahko omenjene enačbe poenostavimo. Neznane vrednosti Q in h bomo našli že, če bodo hkratno zadovoljeni pogoji kontinuitetne in energijske enačbe. Oba pogoja sta zapisala v sistem enačb:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix} * \begin{bmatrix} Q \\ h \end{bmatrix} = \begin{bmatrix} -A_{10}h_0 \\ q \end{bmatrix} \quad (2.14)$$

Zgornji del matrik sovпада z enačbami za energijske izgube, spodnji z kontinuitetnimi enačbami. Matrika A_{11} pa predstavlja matriko razmerja med energijskimi izgubami tlaka in pretoka. Razlagicjo sicer izraža enačba:

$$h_i = K_i Q_i^m \quad (2.15)$$

Parametria m še nismo srečali. Je eksponent pretoka za posamezen model energijskih izgub. Matriko A_{11} zato zapišemo kot:

$$A_{11} = \begin{bmatrix} K_1 |Q_1|^m & & & \\ & K_2 |Q_2|^m & & \\ & & \dots & \\ & & & K_n |Q_n|^m \end{bmatrix} \quad (2.16)$$

Sistem enačb 2.14 je nelinearen in onemogoča direkten izračun rešitve. Po drugi strani je primeren za iterativni numerični izračun. V ta namen sistem 2.14 najprej odvajamo in dobimo:

$$\begin{bmatrix} N A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix} * \begin{bmatrix} \delta Q \\ \delta h \end{bmatrix} = \begin{bmatrix} \delta E \\ \delta q \end{bmatrix} \quad (2.17)$$

V sistemu je N diagonalna matrika z eksponenti m iz enačbe 2.15 in korendenčne matrike 2.16. δQ in δh sta ostanka pri iteracijskem reševanju sistema enačb. Z njima se preverja zaustavitveni pogoj in sta v k -ti iteraciji, definirana kot:

$$\delta E = A_{11} Q^k + A_{12} h^k + A_{10} h \delta q = A_{21} Q^k - q \quad (2.18)$$

Vrednosti iskanih vektorjev Q in h se v skladu z Newton - Raphsenovo metodo iz razdelka 2.4.3 spreminjajo v skladu z enačbo [7]:

$$Q_{k+1} = Q_k + \delta Q \quad (2.19)$$

$$h_{k+1} = h_k + \delta h \quad (2.20)$$

Todini in Pilati [10] tako definirata rekurzivno formulo za izračun h_{k+1} , s pomočjo katere se izračuna Q_{k+1} . Enačbi rekurzivno rešujemo, dokler ostanki 2.18 niso manjši od neke zahtevane tolerance.

$$h_{k+1} = -(A_{21}N^{-1}A_{11}^{-1}A_{12})^{-1}(A_{12}N^{-1}(Q_k + A_{11}^{-1}A_{10}h_0) + (q - A_{21}Q_k)) \quad (2.21)$$

$$Q_{k+1} = (1 - N^{-1})Q_k - N^{-1}A_{11}^{-1}(A_{12}h_{k+1} + A_{10}h_0) \quad (2.22)$$

Časovna analiza gradientne metode

Časovna zahtevnost gradientne metode bo temeljila na analizi zahtevnosti matričnih operacij v rekurzivnih enačbah 2.21 in 2.22. Najprej si pogledimo inverz diagonalne matrike. Ta je numerično enostavnejša operacija od inverza ostalih matrik. Izračunamo ga lahko v $\mathcal{O}(n)$, kot je vidno v enačbi:

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \frac{1}{a_{nn}} \end{bmatrix} \quad (2.23)$$

Časovne zahtevnosti matričnih operacij so tako:

Operacija	Časovna zahtevnost
Inverz diagonalne matrike	$\mathcal{O}(n)$
Množenje diagonalnih matrik	$\mathcal{O}(n)$
Vsota dveh matrik	$\mathcal{O}(n^2)$
Inverz matrike s Strassenovo metodo	$\mathcal{O}(n^{2.807})$
Množenje matrik s Strassenovo metodo	$\mathcal{O}(n^{2.807})$
Množenje $n \times m$ in $n \times p$ matrike	$\mathcal{O}(nmp)$

Velikostni razredi matrik iz formul 2.21 in 2.22 so:

- A_{12} - (ne, nn)
- A_{10} - (ne, no)
- N - (ne, ne)
- A_{11} - (ne, ne)

Pri čemer so:

- nn - Število vozlišč z neznanim tlakom
- no - Število vozlišč z znanim tlakom
- ne - Število cevi z neznanim pretokom

Izpeljavo časovne zahtevnosti smo precej poenostavili brez prevelike izgube natančnosti, če upoštevamo, da je v vodovodnih sistemih število vozlišč z neznanim tlakom in število povezav približno enakega velikostnega razreda, torej $ne \approx nn = n$. Poleg tega lahko postavimo tudi relacijo, da je število vozlišč z neznanim tlakom za nekaj velikostnih razredov večje od vozlišč z znanim tlakom $1 \approx no \ll nn = n$.

Najprej razpišimo časovne zahtevnosti po posameznih členih za 2.21:

- V členu $-(A_{21}N^{-1}A_{11}^{-1}A_{12})^{-1}$ iz enačbe 2.21 imamo tako: 2 inverza diagonalne matrike, množenje diagonalne matrike, dve množenji matrik in inverz matrike. Časovna zahtevnost tega dela je $\mathcal{O}(3 * n^{2.807} + 3 * n)$.

- V členu $(A_{12}N^{-1}(Q_k + A_{11}^{-1}A_{10}h_0) + (q - A_{21}Q_k))$ imamo dva inverza diagonalne matrike, tri množenja $(n \times n)$ z $(n \times 1)$ matriko, dve množenji vektorjev dolžine n ter tri seštevanja $(n \times 1)$ matrik. Vse skupaj seštejemo v časovno zahtevnost $\mathcal{O}(3 * n^2 + 7 * n)$
- Za množenje obeh členov potrebujemo še $\mathcal{O}(n^2)$, kar vse skupaj zneso za polinomsko časovno zahtevnost $\mathcal{O}(3 * n^{2.807} + 7 * n^2 + 10 * n)$. Ker lahko konstante in člene z nižjimi koeficienti okrajšamo, zapišemo zahtevnost enostavneje $\mathcal{O}(n^{2.807})$.

Analiza po členih enačbe 2.22 je nekoliko enostavnejša:

- Člen $(1 - N^{-1})Q_k$ vsebuje en inverz diagonalne matrike, eno seštevanje in množenje matrike z vektorjem. Časovna zahtevnost je $\mathcal{O}(2 * n^2 + n)$
- Člen $N^{-1}A_{11}^{-1}(A_{12}h_{k+1} + A_{10}h_0)$ sestavljajo dva inverza diagonalnih matrik, štiri množenja matrike z vektorjem in seštevanje vektorjev. Zahtevnost tega člena je $\mathcal{O}(4 * n^2 + 3 * n)$.
- Oba člena seštejemo v $\mathcal{O}(6 * n^2 + 4 * n)$ in okrajšamo v $\mathcal{O}(n^2)$.

Zahtevnost enega koraka gradientne metode je tako $\mathcal{O}(n^{2.807})$, ker pa metoda hitro konvergira in najde dovolj dobro rešitev v nekaj iteracijah, se enačba pomnoži le s konstanto, katero lahko zanemarimo.

2.4.5 Reševanje razširjene časovne simulacije

V razdelku 2.4.4 smo obravnavali reševanje hidravličnega modela in njegovih enačb za sisteme s stalnim tokom, v tem pa si pogledajmo iskanje rešitev sistemov s spremenljivimi pogoji. Ker vemo, da poraba vode niha tako dnevno kot sezonsko, se z njo spreminjajo tudi hidravlične razmere v cevovodnem sistemu. V kolikor so spremembe porabe "dovolj" počasne, se njihov vpliv pozna po celotnem sistemu [6] in ga lahko opazujemo na nivojih gladine v vodohranih. Reševanje razširjene časovne simulacije izvajamo v diskretnih časovnih intervalih. V vsakem koraku rešimo sistem enačb, opisan v 2.4.4, in poiščemo neznane pretoke cevi in tlake vozlišč. Poleg tega v vsakem intervalu izvedemo še dodatno prilagoditev, v kateri

se popravijo nivoji vode v vodohranih. V simulaciji s stalnim tokom so bile njihove gladine nespremenljive in so služile kot referenčna točka za reševanje sistema enačb, shranjene v spremenljivki h_0 . Gladine v vodohranih se na vsakem koraku osvežijo v skladu z enačbo [7]:

$$\frac{\delta}{\delta t} \int_{KV} \delta V = \frac{\delta(H_T A_T)}{\delta t} = A_T \frac{\delta H_T}{\delta t} = Q_{t,vtok} - Q_{t,iztok} \quad (2.24)$$

V enačbi so $Q_{t,iztok}$ in $Q_{t,vtok}$ pritoki in iztoki cevovodov, H_T nivo gladine v časovnem koraku T in A_T presek vodohrana. Slednji so ponavadi pravokotnih ali okroglih oblik, njihov presek pa se z višino ne spreminja, zato lahko zgornjo enačbo poenostavimo v:

$$A_T \frac{H_{T,t+\Delta t} - H_{T,t}}{\Delta t} = Q_{t,vtok} - Q_{t,iztok} \quad (2.25)$$

Ko iz enačbe izpostavimo iskani člen $H_{T,t+\Delta t}$, dobimo:

$$H_{T,t+\Delta t} = H_{T,t} + (Q_{t,vtok} - Q_{t,iztok}) \frac{\Delta t}{A_T} \quad (2.26)$$

V enačbi je $H_{T,t+\Delta t}$ iskana gladina na koncu časovnega intervala, $H_{T,t}$ znana gladina na začetku časovnega intervala, $Q_{t,vtok}$ in $Q_{t,iztok}$ pritok in iztok v vodohran in Δt trajanje časovnega intervala.

Določanje dolžine časovnega intervala Δt je v nekaterih pogledih sorodno kvalitativnemu sklepanju (ang. Qualitative reasoning), podpodročju umetne intiligence [19]. Po tem, ko smo našli rešitev za interval T , določimo časovni korak do naslednjega intervala kot minimum:

1. Časa do začetka naslednjega časovnega intervala
2. Najkrajšega časa, dokler se kateri od vodohranov napolni ali izprazni
3. Najkrajšega časa, dokler nivo katerega od vodohranov doseže nivo, ki sproži delovanje katere od črpalk
4. Časa, dokler druga pravila ne povzročijo spremembe v cevovodnem sistemu.

Čase, v katerih nivoji vodohranov dosežejo neko kritično točko, izračunamo na podlagi izračunanih vtokov in iztokov v koraku T . Poenostavlja se, da se nivoji gladin spreminjajo linearno s časom.

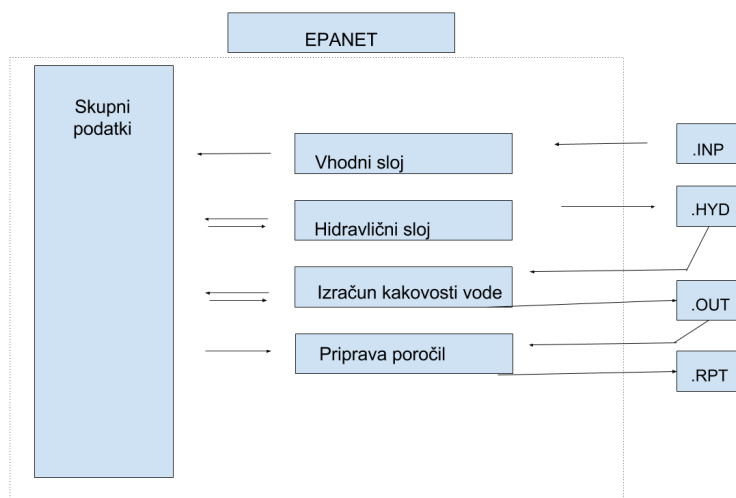
1. Če je bila kritična točka dosežena, preverimo v t.i. intervalu za preverjanje pravil. Slednji je ponavadi nastavljen na $\frac{1}{10}$ dolžine intervala. Če naj bi hidravlični izračun opravili vsaj vsako uro, se pravila preverjajo na vsakih 6 minut.
2. V vsakem od intervalov za preverjanje pravil se določijo novi nivoji gladin vodohranov in popravi ura sistema.
3. Če je bila kritična točka dosežena, vstavimo pravila (naprimer vklop črpalke) v prioritetno vrsto. V kolikor je bilo doseženih več kritičnih točk, se pravila razvrstijo glede na prioriteto.
4. Če prioritetna vrsta s pravili ni prazna, izračunamo hidravlično simulacijo za celoten sistem s popravljeno situacijo, zaradi zadnje spremembe, proces določanja novega časovnega intervala pa se prične znova. Nova iteracija se začne z novimi, pravkar izračunanimi vrednostmi pretokov in tlakov.

2.5 Knjižnica EPANET

EPANET je odprtokodno in prostodostopno orodje za modeliranje vodovodnih sistemov. Razvija ga Ameriška agencija za varovanje okolja (ang. United States Environmental Protection Agency), na kratko EPA. Omogoča tako simulacije sistemov s stalnim tokom kot tudi razširjeno časovno simulacijo hidravličnih razmer v tlačnih sistemih. Namenjeno je tudi raziskovalni sferi, prvič pa je bilo objavljeno leta 1993.

Nadgradnja EPANET 2 je orodje, ki je dostopno v kombinaciji z grafičnim uporabniškim vmesnikom in je v taki obliki najpogostejše uporabljano. Druga oblika uporabe je kot dinamična knjižnjica, v stroki dobro poznana kot "epanet2.dll". Napisana je v jeziku C [6], uporablja pa se kot osnova za izračun hidravličnih razmer v različnih aplikacijah. Do funkcij knjižnjice lahko dostopamo z popularnimi inženirskimi programskimi jeziki, kot so MATLAB, Python in drugi.

Orodje lahko v grobem razdelimo v več slojev, kot je vidno na sliki 2.5:



Slika 2.5: Sestavni deli knjižnjice EPANET 2. [6]

1. Vhodni sloj prejme datoteko s končnico `.inp`, v kateri je opisano tako celotno omrežje kot tudi parametri simulacije. Ta del prebere datoteko, preveri pravilnost zapisa, pretvori podatke v pravilno obliko in jih shrani v skupni pomnilnik.
2. V hidravličnem sloju se izračuna simulacija stalnega toka ali razširjena časovna simulacija. Rezultati vsakega časovnega intervala se zapisujejo v dvojiško datoteko `.hyd`. Nekateri od teh intervalov predstavljajo vmesne intervale, ki niso specificirani s strani uporabnika, ampak so bili izračunani, ker je v sistemu dosežena neka kritična točka, kot je opisano v razdelku 2.4.5.
3. Sloj za izračun kakovosti vode po potrebi pridobi podatke iz sloja za hidravliko.
4. Po potrebi se rezultati zapišejo v datoteko s poročilom `.rpt`, v kateri najdemo tudi podatke o morebitnih napakah in opozorilih.

S pomočjo funkcij lahko uporabnik dostopa do posameznih elementov vodovodnega sistema in njihovih lastnosti, kar je še posebej pomembno v nadaljevanju. Možni sta obe funkciji, torej branje (ang. *Get*) in zapis nove vrednosti (ang. *Set*). Enostaven primer takšnih klicev je viden v spodnji kodi, kjer se klici knjižnjice izvajajo v programskem jeziku Python 2.7.

```
et.ENopenH()
et.ENinitH(0)
while True :
    # Izvedemo simulacijo za naslednji interval
    ret , t=et.ENrunH()
    time.append(t)
    # Beremo tlake v vozliscih
    for i in range(0,len(nodes)):
        ret , p=et.ENgetnodevalue(i+1, et.EN.PRESSURE )
        pres[i].append(p)
    ret , tstep=et.ENnextH()
    if (tstep <=0):
        break
ret=et.ENcloseH()
```


Poglavje 3

Genetski algoritmi

3.1 Metodologija

Evolucijski algoritmi (v nadaljevanju EA) so metahevristični optimizacijski algoritmi, podpodročje umetne inteligence in orodje za preiskovanje prostora. Uporabljajo mehanizme, ki oponašajo naravne procese evolucije, kot so reprodukcija, mutacije in naravna selekcija [11]. Za evolucijske algoritme je značilno, da rešitev ne iščejo nad eno samo instanco problema, temveč optimizirajo množico potencialnih rešitev, ki ji v nadaljevanju rečemo **populacija**, posamezni rešitvi pa **osebki**. Že po terminologiji vidimo vzporednice z biologijo. Podobno kot v naravi, se tudi v evolucijskih algoritmi osebki ustvarijo, živijo neko življenjsko dobo in odmrejo. Njihova življenjska doba je odvisna od kvalitete osebkov, slednjo v EA predstavljajo t.i. **ciljne funkcije** (ang. fitness functions). Populacija se razvija tekom evolucije, v kateri boljši osebki preživijo dlje, slabši pa ponavadi odmrejo.

Evolucijski algoritmi so široko uporabljeno optimizacijsko orodje na področjih, kot so ekonomija, marketing, kemija, robotika, fizika in strojništvo. Dobro se obnesejo pri iskanju "rešitev izven okvirjev". Pri dovolj velikih populacijah se ne zataknejo pri lokalnih ekstremih ampak praviloma najdejo rešitev blizu globalnega optimuma.

Slaba lastnost EA je ponavadi časovna potratnost, ki je večinoma povezana s časom računanja ciljnih funkcij. Kot je vidno v nadaljevanju, so lahko ciljne funkcije zelo kompleksne. Po drugi strani EA pogosto učinkovito rešijo kompleksne

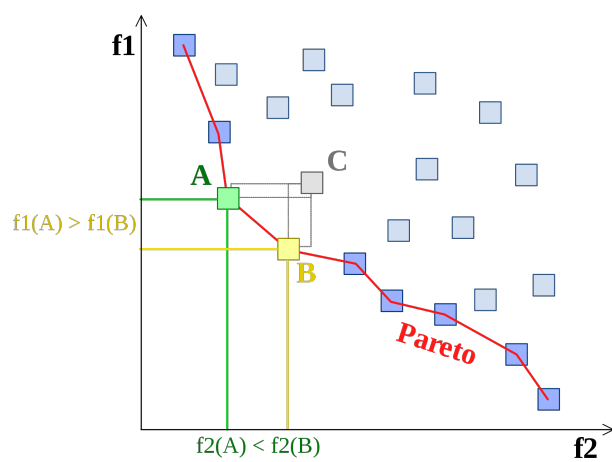
probleme brez močne povezave med časovno zahtevnostjo in velikostjo problema.

Genetski algoritmi so se kot podpodročje evolucijskih algoritmov pojavili v šestdesetih letih prejšnjega stoletja na Univerzi v Michiganu. Skupina pod vodstvom Johna Hollanda se je sprva ukvarjala z evolucijskimi algoritmi kot splošnim načinom reševanja problemov v računalništvu, ne pa iskanju konkretne rešitve posameznega problema. V knjigi *Adaptation in Natural and Artificial Systems* so predstavili genetske algoritme kot abstrakcijo biološke evolucije. Člane populacije (osebke) so zakodirali v gen, predstavljen kot niz bitov, nad katerimi so izvajali operacije mutacije, križanja in inverza [13]. Omenjene manipulacije genov smo podrobneje opisali v naslednjih razdelkih. Prehod med dvema populacijama so izvedli na podlagi "naravne selekcije", kjer so boljši osebki preživeli, slabši pa ne.

Do danes so raziskovalci predlagali veliko različnih GA, ki se med seboj razlikujejo predvsem po načinu manipulacije s populacijo, gradnji novih osebkov in selekciji starih. Za kodiranje gena se uporabljajo kompleksnejše podatkovne strukture; od tabel celih in decimalnih števil do "po meri"narejenih objektov. Ker se GA pogosto uporabljajo kot orodje za reševanje optimizacijskih problemov, so se razvili algoritmi za učinkovito reševanje dvo in večkriterijskih optimizacijskih problemov. Slednji ocenjujejo kvaliteto osebka na podlagi več ciljnih funkcij, ki so si med seboj pogosto nasprotujoče. Rešitev takšnega problema je podmnožica populacije, kjer so osebki optimalni po vsaj enem kriteriju. V nadaljevanju smo opisali enega od takšnih algoritmov.

3.2 Pareto optimalne rešitve

Kot je bilo bežno omenjeno v prejšnjem razdelku, se v praksi srečujemo z optimizacijo po več različnih kriterijih. Vzemimo enostaven primer nakupa avtomobila, pri katerem se odločamo na podlagi dveh parametrov: letnik izdelave in cena. Odločamo se torej med novejšimi in dražjimi ali starejšimi in cenejšimi avtomobili. Hitro opazimo, da sta si kriterija cena in starost med seboj konfliktna. Rešitev takšnega optimizacijskega problema je množica Pareto-optimalnih rešitev. V konkretnem primeru gre za množico avtomobilov, ki imajo pri dani starosti najnižjo ceno. Rešitve takšnega problema se grafično lepo predstavijo kot **Pareto fronta**, kot je razvidno na sliki 3.1.



Slika 3.1: Pareto-fronta povezuje množico optimalnih rešitev [12].

3.2.1 Nedominantno sortiranje

Algorithm 1 fast nondominant sort

```

for  $p \in P$  do // pri vseh parih preverimo
  for  $q \in P$  do
    if  $p < q$  then // ce q dominira p
       $S_p = S_p \cup q$  // q dodamo v mnozico S - resitve ki dominirajo p
    else if  $p > q$  then // ce p dominira q
       $n_p = n_p + 1$  // povecamo n - stevilo resitev, ki jih dominira p
    end if
  end for
  if  $n_p = 0$  then // najboljse resitve
     $F_1 = F_1 \cup p$ 
  end if
end for

 $i = 1$ 
while  $F_1 \neq \emptyset$  do
   $H = \emptyset$ 
  for  $p \in F_i$  do // za vse resitve p v trenutni Pareto-fronti
    for  $q \in S_p$  do // preverimo resitve q, ki jih p dominira
       $n_q = n_q - 1$ 
      if  $n_q = 0$  then // ce so bile dominirane samo s strani p
         $H = H \cup q$  // jih dodamo v drugo Pareto-fronto
      end if
    end for
  end for
   $i = i + 1$  // povecamo rang
   $F_i = H$  // elementi druge Pareto-fronte so sedaj v trenutni
end while

```

Pravkar smo spoznali koncept "Pareto-optimalnih" rešitev, ki so, kot že ime pove, vse optimalne. Pripisimo tem rešitvam neko lastnost **rang** z vrednostjo 1, saj so

najboljše. Če slednje začasno odstranimo in zgradimo novo Pareto-fronto, dobimo množico osebkov z rangom 2. Postopek lahko tako nadaljujemo in tako dobimo učinkovit enodimenzionalni kriterij za kvaliteto rešitev. Razvrščanje osebkov na ta način imenujemo nedominantno sortiranje (ang. non dominant sorting), relaciji dveh osebkov pa **Pareto-dominacija**.

Najprej moramo definirati pojem dominantnosti. Za osebk a in b velja, da je a dominanten b (v nadaljevanju $a > b$, če za vsaj eno lastnost m velja, da je $a_m > b_m$). Iz definicije je razvidno, da je dominantnost lahko obojestranska in se lahko zgodi $a > b$ in hkrati $a < b$. To velja za vsaki dve rešitvi iz Pareto-fronte.

Na podlagi omenjene relacije urejenosti deluje algoritem (glej algoritem 1) za nedominantno sortiranje [14]. Najprej za vsako rešitev določimo:

- n_i - število rešitev, ki dominirajo rešitev i
- S_i - množico rešitev, ki jih dominira rešitev i

Ta del je enostavno implementiran tako, da pregleda vse možne pare rešitev, za kar potrebujemo $\mathcal{O}(N^2)$ časa, pri čemer je N število vseh rešitev. Ker pri vsaki primerjavi pregledamo vse kriterije, katerih je m , je natančnejša časovna zahtevnost tega dela $\mathcal{O}(m * N^2)$.

V nadaljevanju pregledamo vse rešitve in izberemo tiste z $n_i = 0$. Te sestavljajo prvo Pareto-fronto F_1 , saj jih ne dominira nobena druga rešitev. Poleg tega pri vsaki rešitvi iz Pareto-fronte pregledamo njeno množico S_i . Pri vsakem $q \in S_i$ zmanjšamo n_q . Če n_q doseže 0, ga dodamo v H , naslednjo Pareto fronto. S tem dosežemo, da so v F_1 rešitve z rangom 1, v množici H , ki bo postala F_2 , pa rešitve, ki so dominirane samo s strani elementov F_1 in imajo zato rang 2. Ko smo našli vse rešitve z rangom 2, se zanka premakne na iskanje rešitev z rangom 3 in tako naprej. Vseh možnih front je lahko največ N , vsak element pa lahko dominira največ N ostalih elementov. Zato je časovna zahtevnost tega dela algoritma $\mathcal{O}(N^2)$.

Časovna zahtevnost celotnega nedominantnega sortiranja je $\mathcal{O}(N^2) + \mathcal{O}(m * N^2)$, kar je $\mathcal{O}(m * N^2)$. Ker lahko vsaka rešitev dominira največ N ostalih, katere hranimo v S_i , vseh rešitev pa je N je prostorska zahtevnost nedominantnega sortiranja $\mathcal{O}(N^2)$ [14].

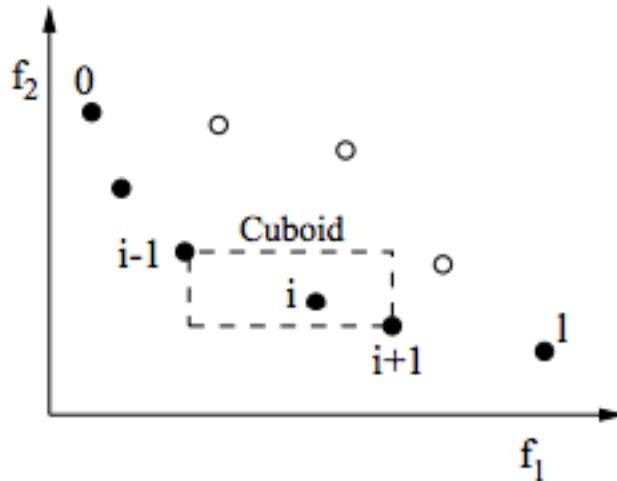
3.2.2 Elitizem

Preden opišemo jedro algoritma NSGA-2, opišimo še pojem elitizma. Gre za izraz iz sociologije, ki opisuje množico posameznikov (elito), ki imajo določeno lastnost in imajo zaradi nje večji vpliv v družbi ter določene imunitete. Izraz v kontekstu GA določa, da se tekom selekcije ohrani množica najboljših rešitev in prenese v naslednjo generacijo.

3.2.3 Metrika zgoščenosti

Za primerjavo dveh rešitev istega ranga potrebujemo še dodatno metriko - metriko zgoščenosti. Ta nam za rešitev i opiše zgoščenost rešitev z istim rangom v njeni okolici. Uporabnost metrike je razvidna v naslednjem razdelku, zato jo tukaj samo opišemo.

Pri i ti rešitvi izračunamo povprečje razdalj do dveh točk na vsaki strani rešitve i pri vsaki od ciljnih funkcij, kot je vidno na sliki 3.2. Tako dobimo vrednost i_{dist} , ki ponazarja velikost stranice največje superkocke okrog rešitve i , ki ne vključuje nobene druge rešitve.



Slika 3.2: Superkocka okrog rešitve i [14].

Algoritem za izračun metrike zgoščenosti je podan z 2 [14]:

Algorithm 2 Crowding distance calculation

```

 $l = |I|$  ▷ Število rešitev z istim rangom
for  $i$  do
     $i_{dist} = 0$ 
end for
for objective  $m$  do
     $I = \text{sort}(I, m)$ 
     $I[0] = \text{inf}$  // prvi in zadnji element imata neskončno superkocko
     $I[l] = \text{inf}$ 
    for  $j = 1$  to  $l - 1$  do // pri ostalih je stranica enaka razdalji do sosedov
         $I_{dist} = I_{dist} + (I_m[j + 1] - I_m[j - 1])$ 
    end for
end for

```

pri tem je $I_m[j]$ vrednost ciljne funkcije m rešitve j .

Ocena časovne zahtevnosti tega algoritma je enostavna. Najpotratnejša operacija je urejanje tabele rešitev dolžine l , ki se zgodi m krat - enkrat za vsako ciljno funkcijo. Tako je časovna zahtevnost $\mathcal{O}(m * l \log l)$.

3.2.4 Primerjava dveh rešitev

Ko smo definirali nedominantno razvrščanje in metriko zgoščenosti, lahko definiramo relacijo urejenosti med dvema rešitvama a in b [7].

$$a \geq b \longrightarrow (a_{rang} > b_{rang}) \vee ((a_{rang} = b_{rang}) \wedge (a_{dist} > b_{dist})) \quad (3.1)$$

Relacija v praksi pomeni, da rešitvi najprej primerjamo po rangi in sicer tako, da je boljša rešitev tista z nižjim rangom. Če sta rešitvi istega ranga, ju primerjamo glede na metriko zgoščenosti in sicer za boljšo rešitev vzamemo tisto z večjo i_{dist} . Na ta način zagotovimo, da se kot boljše izberejo rešitve, ki imajo bolj oddaljene sosede. Tako dosežemo večjo razpršenost, ki je ključnega pomena za učinkovito delovanje algoritma.

3.3 Delovanje NSGA-2

Algoritem NSGA-2 (ang. non-dominated sorting genetic algorithm 2) uporablja nedominantno sortiranje pri selekciji populacije. Gre za elitistični genetski algoritem.

Ob inicializaciji se najprej ustvari prva populacija P_0 velikosti N in množica otrok Q_0 . Od tu naprej se algoritem izvaja znotraj zanke, kjer vsaka ponovitev ponazarja eno generacijo (glej algoritem 3) [14]:

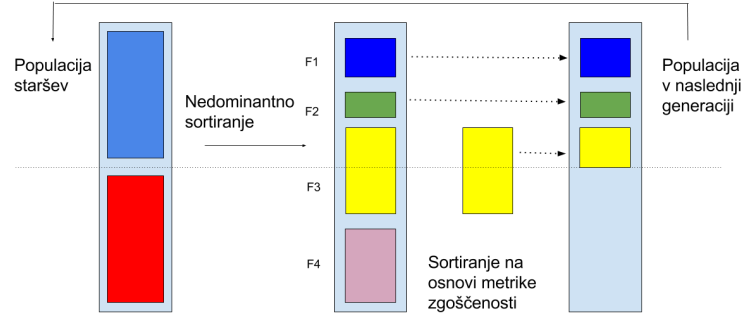
Algorithm 3 NSGA-2

```

while ! stopping criteria satisfied do
     $R_t = P_t \cup Q_t$ 
     $F = \text{fast nondominant sort } (R_t) // \text{resitve urejene po rangih}$ 
    while  $|P_{t+1}| < N$  do // do zapolnitve populacije
        crowding distance calculation ( $F_i$ )
         $P_{t+1} = P_{t+1} \cup F_i // \text{dodamo resitve ranga } i$ 
    end while
     $\text{sort}(P_{t+1})$ 
     $P_{t+1} = P_{t+1}[0 : N] // \text{odrezemo odvecne resitve}$ 
     $Q_{t+1} = \text{create children } (P_{t+1}) // \text{ustvarimo nove otroke}$ 
     $t = t + 1$ 
end while

```

Ob vsaki iteraciji najprej združimo množici staršev (P_t) in otrok (Q_t) v skupno množico R_t . Rešitve znotraj R_t uredimo s pomočjo nedominantnega sortiranja. Tako vsaki rešitvi pripišemo ustrezen rang. Novo populacijo (P_{t+1}) generiramo tako, da najprej vstavljamo rešitve, najprej tiste z nižjim rangom nato pa višje, dokler velikost nove populacije ne preseže dovoljene velikosti N . Ob vsakem vstavljanju rešitvam določimo metriko zgoščenosti. V tem trenutku se ponavadi zgodi, da imamo v P_{t+1} več elementov kot N , ker je število rešitev z istim rangom večje kot število prostih mest v P . V tem koraku si pomagamo z metriko zgoščenosti in relacijo urejenosti, opisanima v prejšnjih razdelkih. Elemente znotraj populacije P_{t+1} uredimo in v naslednji krog prenesemo samo N najboljših. Tako generiramo populacijo P_{t+1} velikosti N , kot je lepo razvidno na sliki 3.3.



Slika 3.3: Selekcija najboljših rešitev na podlagi nedominantnega sortiranja in metrike zgoščevanja.

V nadaljevanju se iz populacije P_{t+1} generirajo otroci s pomočjo mutacije in križanja. Pri izbiri staršev se uporablja sistem binarnega turnirja, ki smo ga opisali v nadaljevanju.

Denimo, da v tem trenutku ne upoštevamo časovne zahtevnosti evalvacijske funkcije, ki je odvisna od tipa problema in jo bomo podrobneje opisali v nadaljevanju. Zahtevnost algoritma je tako odvisna od osnovnih operacij z zahtevnostmi v tabeli 3.1 in je tako enaka $\mathcal{O}(m * N^2)$.

Operacija	Časovna zahtevnost
Nedominantno sortiranje	$\mathcal{O}(m * N^2)$
Določitev metrike zgoščevanja	$\mathcal{O}(m * l \log l)$
Urejanje populacije za binarni turnir	$\mathcal{O}(2N \log 2N)$

Tabela 3.1: Časovne zahtevnosti operacij znotraj algoritma NSGA-2

3.4 Genetski operatorji

3.4.1 Reprodukcijska

Reprodukcijska je ponavadi prva operacija, ki jo izvedemo nad populacijo [17]. Iz populacije izbere starše, ki jih v nadaljevanju uporabimo pri križanju. V nadaljevanju si pogledamo dva načina izbire, pri kateri želimo upoštevati idejo evolucije in boljšim dati več možnosti za reprodukcijo.

1. Izbira s pomočjo ruletnega kolesa

Izbira staršev s pomočjo ruletnega kolesa je enostaven postopek, kjer je verjetnost izbire starša iz populacije proporcionalna njegovi kvaliteti. Ideja je, da imajo boljši osebki večjo možnost, da postanejo starši. Po drugi strani pa imajo tudi najslabši v populaciji možnost, le da je ta manjša, saj se lahko zgodi, da imajo tudi slabši starši kakšne pomembne lastnosti za razvoj populacije.

Primer izračuna verjetnosti izbire je prikazan v tabeli 3.2.

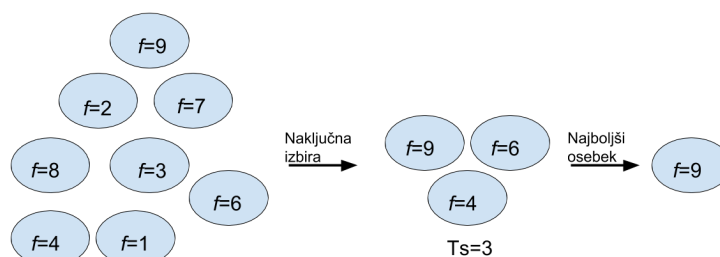
	Kvaliteta	P_i	$N * P_i$
Osebek 1	169	0.14	0.58
Osebek 2	576	0.49	1.97
Osebek 3	64	0.06	0.22
Osebek 4	361	0.31	1.23
Vsota	1170	1.00	4.00
Povprečje	293	0.25	1.00
Maksimum	576	0.49	1.97

Tabela 3.2: Populacija petih osebkov.

2. Turnirska izbira

Sistem turnirske izbire je najpogostejši način reprodukcije v GA [16]. Je časovno učinkovita in enostavna za implementacijo. Iz populacije najprej naključno izberemo n posameznikov. Slednji med seboj tekmujejo; izbran je

najkvalitetnejši posameznik. Število izbranih osebkov n je pogosto 2; takemu sistemu rečemo binarni turnir (ang. binary tournament). Večje velikosti turnirjev n pomenijo upad diverzitete izbrane populacije, saj manjši delež populacije prispeva genski material [15].



Slika 3.4: Turnirska izbira z velikostjo $n = 3$.

Sistem turnirske izbire ima več prednosti: omogoča tudi slabšim posameznikom možnosti izbire, zagotavlja, da dominantni osebki ne prevladajo, je enostaven za paralelizacijo in ne potrebuje normalizacije kriterijske funkcije.

3.4.2 Križanje

Križanje je genetski operator, ki iz dveh staršev, predstavljenih z dvema genom, ustvari potomce. Ideja je, da so lahko rešitve zakodirane v genome potomcev boljše od staršev, v kolikor bi pri križanju pobrali boljše lastnosti obeh staršev. Križanje se izvaja z določeno verjetnostjo, ki je definirana s strani uporabnika. V nadaljevanju na hitro pregledamo nekaj najpogostejših metod omenjene operacije.

1. Enotočkovno križanje

Enotočkovni operator naključno izbere točko križanja - indeks v tabeli. Do omenjenega indeksa kopira vrednosti prvega starša, naprej pa od drugega in tako ustvari prvega otroka. Ker smo na ta način mogoče pobrali samo slabe

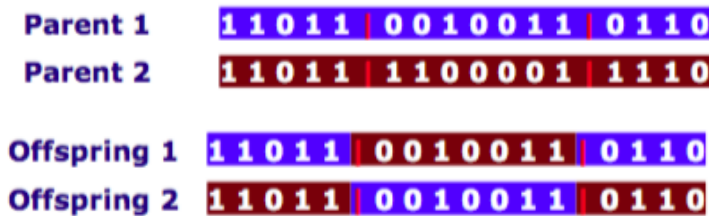
lastnosti obeh staršev, se drugi otrok ustvari na ravno obraten način, kot je razvidno na sliki 3.5:



Slika 3.5: Enotočkovno križanje [17].

2. Dvotočkovno križanje

Dvotočkovno križanje je podobno enotočkovnemu, le da tu izberemo dva indeksa znotraj tabele, kot je vidno na sliki 3.6.



Slika 3.6: Dvotočkovno križanje [17].

3. Uniformno križanje

Uniformni operator ne uporablja sekcij genoma staršev za generiranje otrok, ampak se za vsak gen posebej odloči, od katerega starša bo. Izbira gena je pogojena z verjetnostjo mešanja. Če je ta 0.5 se pričakuje, da bo približno polovica genov od prvega, polovica pa od drugega starša, kot je to vidno na sliki 3.7.

Parent 1	1	1	0	1	1	0	0	1	0	0	1	1	0	1	1	0
Parent 2	1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	0
Offspring 1	1 ₁	1 ₂	0 ₂	1 ₁	1 ₁	1 ₂	1 ₂	0 ₂	0 ₁	0 ₁	0 ₂	1 ₁	1 ₂	1 ₁	1 ₁	0 ₂
Offspring 2	1 ₂	1 ₁	0 ₁	1 ₂	1 ₂	0 ₁	0 ₁	1 ₁	0 ₂	0 ₂	1 ₁	1 ₂	0 ₁	1 ₂	1 ₂	0 ₁

Slika 3.7: Uniformno križanje [17].

3.4.3 Mutacija

Mutacija je operacija, ki se izvaja nad genomom enega samega osebk. Ponavadi gre za majhne spremembe genoma na enem ali dveh mestih. S tem dosežemo, da se genetska raznolikost spreminja iz generacije v generacijo. Na ta način lahko algoritem preišče širši prostor in najde rešitve, ki bi bile samo z križanjem osebkov nedosegljive.

Vzemimo za primer trivialno populacijo petih osebkov, katerih genom je sestavljen iz binarnih vrednosti, prikazan v tabeli 3.3.

	G_1	G_2	G_3	G_4	G_5
Osebek 1	1	0	1	0	1
Osebek 2	0	1	1	0	1
Osebek 3	0	0	1	1	1
Osebek 4	1	0	0	1	1
Osebek 5	1	1	0	1	1

Tabela 3.3: Populacija petih osebkov.

Neglede na tip križanja, ki ga izberemo in kombinacijo staršev, ne moremo doseči otrok z vrednostjo 0 na mestu G_5 . V dvovrednostnem prostoru se pri dovolj veliki populaciji skoraj ne more zgoditi omenjeni scenarij, pri celoštevilskih in decimalnih vrednostih genoma pa je nemogoče zavzeti vse možne vrednosti v začetni populaciji. Mutacija z majhnimi spremembami genoma to omogoča. Poglejmo si nekaj načinov mutacije in tipe kodiranja genomov, kjer so uporabljeni:

1. **Inverzni bit** Mutacija z inverzom bita je najenostavnejši tip mutacije. Uporabna je, kjer so geni predstavljeni v dvovrednostnem sistemu. Izbranemu genu znotraj genoma samo zamenjamo vrednost z 1 na 0 ali obratno. Operacijo lahko izvedemo na enem ali več genih, kot je vidno na sliki 3.8.

Original offspring 1	1 1 0 1 1 1 1 0 0 0 0 1 1 1 1 0
Original offspring 2	1 1 0 1 1 0 0 1 0 0 1 1 0 1 1 0
Mutated offspring 1	1 1 0 0 1 1 1 0 0 0 0 1 1 1 1 0
Mutated offspring 2	1 1 0 1 1 0 1 1 0 0 1 1 0 1 0 0

Slika 3.8: Primer mutacije z inverzom bita [17].

2. **Mejna mutacija** Mejna mutacija je primerna za gene, predstavljene s celimi ali decimalnimi števili. Izbrani gen zamenja z zgornjo ali spodnjo dovoljeno mejo.
3. **Uniformna mutacija** Pri uniformni mutaciji se gen zamenja z naključno vrednostjo med zgornjo in spodnjo mejo. Tudi ta mutacija je smiselna na celoštevilskih in decimalnih genih.
4. **Gaussova mutacija** Gaussova mutacija je nekoliko kompleksnejša od prej opisanih, a zelo dobro opravi svojo nalogo predvsem pri genih, kodiranih z decimalnimi števili. Pred operacijo mutacije se najprej opravi statistična analiza populacije in sicer se za množico vseh osebkov izračuna povprečje μ in varianca σ^2 za vsak gen posebej. Nova vrednost gena x se nato izračuna na podlagi omenjenih parametrov z verjetnostjo:

$$f(x \mid \sigma^2, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (3.2)$$

Enostavnejše implementacije Gaussove mutacije ne uporabljajo predhodne statistične analize, temveč uporabljajo vnaprej določene parametre μ in σ^2 .

Poglavje 4

Umerjanje parametrov vodovodnih sistemov

4.1 Uvod

Hidravlični simulacijski modeli vodovodnih omrežij so široko uporabljeno orodje, tako med načrtovanjem kot med obratovanjem in vzdrževanjem. Danes se hidravlične modele relativno hitro vzpostavi na podlagi podatkov, shranjenih v geografsko informacijskih sistemih (ang. Geographic information system), na kratko GIS, ali programih za računalniško podprto načrtovanje (ang. Computer-aided design) na kratko CAD. Kot pri vseh matematičnih modelih potrebujejo tudi parametri hidravličnih simulacijskih modelov umerjanje, preden nam bodo vrnili uporabne rezultate in jih bomo lahko praktično uporabili. Danes je na voljo več tehnik za avtomatsko umerjanje parametrov hidravličnih modelov, tehnologija pa je manj razvita za modele, ki simulirajo kakovost vode. Poleg kalibracijskih orodij potrebujemo tudi standarde preverjanja rezultatov, ki nam povedo, ali je bila kalibracija opravljena v zadostni natančnosti.

V tej nalogi se osredotočamo na kalibracijo hidravličnih modelov. Pri tem skušamo določiti neznane fizikalne karakteristike vodovodnega sistema. Parametri, ki opisujejo te karakteristike, naj bi ob vključitvi v hidravlični model povzročili čim bolj podobno hidravlično situacijo v simulaciji, kot jo opazujemo v naravi. Natančneje iščemo takšne neznane parametre, da bo razlika med meritvami na terenu

in modelnimi napovedmi čim manjša. Pri umerjanju hidravličnih modelov lahko iščemo vrednosti več tipov neznanih parametrov. Najpogosteje pa tekom umerjanja iščemo le vrednosti za t.i. hrapavost, katere učinek smo opisali v razdelku 2.3.3. Omenjeni način umerjanja je uporabljen tudi v tej nalogi.

V tem poglavju so najprej opisane tehnike umerjanja parametrov vodovodnih sistemov. Podrobneje je pregledana tehnika umerjanja na osnovi genetskih algoritmov. Na koncu so opisani še nekateri standardi za verifikacijo kakovosti umerjanja.

4.1.1 Pregled tehnik umerjanja parametrov

Tehnike umerjanja hidravličnih modelov lahko razvrstimo v tri kategorije, ki se med seboj razlikujejo tako po ideji kot po učinkovitosti in uporabnosti za implementacijo. Vse tri na kratko opišemo in tako s primerjavo utemeljimo izbiro najprimernejše za kontekst te naloge:

1. Iterativna metoda

Iterativna metoda deluje na podlagi poizkušanja in napak. Je ročna metoda in jo opravlja inženir. V tem kalibracijskem procesu iterativno poravljamo neznane parametre. V vsakem koraku izračunamo sistem enačb iz razdelka 2.4 in tako pridobimo podatke za pretoke cevi in tlake v vozliščih. Pri tej tehniki se ponavadi uporabi poenastavitev cevovodnega omrežja s pomočjo skeletonizacije. Število parametrov umerjanja se s tem zmanjša. Ta tehnika reši majhne kalibracijske probleme, z majhnim številom parametrov. Slabost te tehnike je počasna konvergenca [2]. Iz te tehnike izhajajo nekatere smernice, ki jih učinkovito uporabljajo bolj napredne metode, opisane v nadaljevanju.

2. Eksplicitne tehnike

Pri eksplicitni kalibraciji rešujemo razširjeni sistem enačb, ki opisuje hidravlične razmere v sistemu s stalnim tokom. Ta razširjeni sistem vsebuje enačbe, ki opisujejo sistem s stalnim tokom in vrsto dodatnih enačb. Slednje izhajajo iz podanih podatkov o izmerjenih tlakih v vozliščih in pretokih v ceveh. Za vsako meritev se doda po ena enačba. Razširjeni sistem enačb se rešuje eksplicitno z Newton-Raphsenovo metodo [2].

Ker je v vsakem sistemu enačb število neznank enako številu enačb, je v eksplicitni metodi kalibracije število neznanih parametrov enako številu meritev. V primeru večjega števila neznank se slednje grupirajo v skupine.

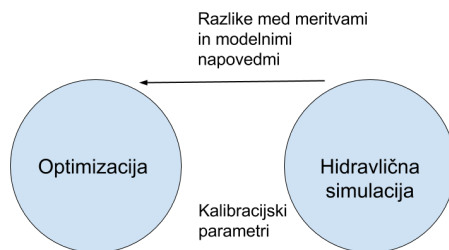
Tovrstne metode imajo precej pomankljivosti [2]:

- (a) Število kalibracijskih parametrov mora biti enako številu meritev.
- (b) Napake pri meritvah se ne upoštevajo. Model predvideva, da so meritve 100% natančne.
- (c) Iz modela je nemogoče izluščiti negotovost modelnih napovedi.
- (d) Metoda potrebuje zahtevna računalniška orodja.

Zaradi zgoraj navedenih razlogov je metoda zgolj teoretično uporabna. V praksi se jo skorajda ne uporablja.

3. Implicitne metode

Med implicitne metode kalibracije štejemo metode, sestavljene iz dveh delov: optimizacijske metode in sistema enačb, ki iščejo neznanke hidravličnega modela. Proces optimizacije sestavljata dva koraka, kot je prikazano na sliki 4.1.



Slika 4.1: Koraki implicitne metode.

Optimizacijska metoda hidravličnemu modelu najprej nastavi neznane parametre, ta pa po izračunu sistema enačb vrne modelne napovedi, torej tlake in pretoke. Metoda minimizira razliko med meritvami in modelnimi napovedmi. Minimizacijsko funkcijo lahko v splošnem zapišemo kot:

$$\min f(x) = \sum_{i=1}^n w_i |\tilde{y}_i - y_i(x)|^p \quad (4.1)$$

V enačbi so \tilde{y}_i izmerjene vrednosti, $y_i(x)$ modelne napovedi ob parameetrih, zapisanih v vektorju x in w obtežni faktorji. Ciljna funkcija torej išče parametre x , da bo odstopanje med meritvami in modelnimi napovedmi najmanjše. Eksponent p je ponavadi nastavljen na 2. V implicitnih metodah lahko vrednosti x dosežejo katerokoli vrednost med minimalno in maksimalno dovoljeno vrednostjo za določen parameter.

Z uporabo implicitne metode se izognemo problemom iz eksplicitnega pristopa k umerjanju parametrov vodovodnega sistema, zato je implicitni pristop uporabljen v sklopu te naloge.

4.1.2 Pregled optimizacijskih pristopov

Pri umerjanju parametrov vodovodnega sistema z implicitno metodo rešujemo, poleg sistema hidravličnih enačb, še večkriterijski optimizacijski problem. Literatura navaja dva osnovna načina optimizacije pri umerjanju.

Nevolucijski pristopi temeljijo na iterativnem procesu optimizacije, kjer se na vsakem koraku izračuna gradient ciljne funkcije v odvisnosti od vseh optimizacijskih parametrov. V ta namen se izračuna Jacobijeva (prvi odvod) in Hessejeva matrika (drugi odvod). Izračuna sta numerično zahtevna, kot je razvidno v nadaljevanju.

Evolucijski pristopi so po drugi strani manj zahtevni in imajo vrsto prednosti:

1. Enostavnejši so za implementacijo in manj računsko zahtevni, saj ne potrebujejo zahtevnih matematičnih in matričnih kalkulacij.
2. Uporabni so pri velikih kalibracijskih sistemih, takih, ki so praktično uporabni.
3. Enostavno jih razširimo z dodatno ciljno funkcijo in dodatnimi meritvami.
4. Enostavnejši so za paralelizacijo.

Zaradi naštetih razlogov je bil v tej nalogi izbran evolucijski pristop za reševanje optimizacijskega problema v implicitni metodi umerjanja parametrov. Uporabili smo večkriterijsko optimizacijo z genetskimi algoritmi s pomočjo algoritma NSGA-II, opisano v razdelku 3.3.

4.2 Umerjanje z genetskimi algoritmi

Utemeljili smo izbiro implicitnega pristopa k umerjanju in natančneje izbiro genetskih algoritmov z NSGA-II v ta namen. Izračun hidravlične simulacije smo opisali v razdelku 2.4, implementacijo NSGA-II pa v razdelku 3.3. V tem razdelku se osredotočimo na združitev obeh orodij v kalibracijski algoritem, ki je bil implementiran v sklopu te naloge.

4.2.1 Oblikovanje ciljnih funkcij

NSGA-II potrebuje za delovanje vsaj dve ciljni funkciji (ang. Objective functions) za ocenjevanje kvalitete posameznika. Slednje vračajo odstopanje med modelnimi napovedmi in meritvami. V sklopu te naloge sta bili uporabljeni dve ciljni funkciji, lahko pa uporabimo tudi tri, kot je razvidno v nadaljevanju.

1. Razlika tlakov

Prva ciljna funkcija deluje na podlagi razlike tlakov. Kot že ime pove, vrača razliko med modelnimi napovedmi pritiskov v vozliščih $H(x \mid \Theta)$ in meritvami tlakov $\widetilde{H}_{i,j}$. Razlika se izračuna za vse časovne intervale $j = 1, \dots, N_{timestamp}$, pri vsakem časovnem intervalu pa se izračuna za vsa merilna mesta $i = 1, \dots, N_{pressure}$. Ciljna funkcija razlike tlakov najprej kvadrira, nato pa sešteje.

$$f(\Theta) = \sum_{i=1}^{N_{pressure}} \left(\sum_{j=1}^{N_{timestamp}} (\widetilde{H}_{i,j} - H(x \mid \Theta))^2 \right) \quad (4.2)$$

Ciljna funkcija vrača povprečno vrednost kvadratne napake (ang. mean square error):

$$of_{pressure} = \frac{f(\Theta)}{N_{pressure} * N_{timestamp}} \quad (4.3)$$

V primeru uporabe utežnih koeficientov bi se vsota (Θ) glasila:

$$f(\Theta) = \sum_{i=1}^{N_{pressure}} \sum_{j=1}^{N_{timestamp}} [\widetilde{H}_{i,j} - H(x \mid \Theta)]^T * w_{i,j} * [\widetilde{H}_{i,j} - H(x \mid \Theta)] \quad (4.4)$$

pri čemer so $w_{i,j}$ utežni koeficienti.

V primeru kalibracije parametrov sistema s stalnim tokom je uporabljen samo en časovni interval, tako je $j = 1$.

2. Razlika pretokov

Ciljna funkcija razlike pretokov je podobna prej opisani ciljni funkciji razlike tlakov. Pri vsakem od časovnih intervalov $j = 1, \dots, N_{timestamp}$ se pri vsakem od merilnih mest $i = 1, \dots, N_{flow}$ izračuna razlika med izmerjeno vrednostjo pretoka $\widetilde{Q}_{i,j}$ in modelno napovedjo $Q(x | \Theta)$. Razlike se najprej kvadrirajo, nato pa povprečijo.

$$of_{flow} = \frac{\sum_{i=1}^{N_{flow}} (\sum_{j=1}^{N_{timestamp}} (\widetilde{Q}_{i,j} - Q(x | \Theta))^2)}{N_{flow} * N_{timestamp}} \quad (4.5)$$

3. Razlika gladin

V nekaterih primerih umerjanja VS se uporablja tretja ciljna funkcija. Slednja uporablja kot kriterij razliko gladin v vodohranih. Pri vsakem od časovnih intervalov $j = 1, \dots, N_{timestamp}$ se pri vsakem od merilnih mest gladine $i = 1, \dots, N_{surface}$ izračuna razlika med izmerjeno vrednostjo gladine $\widetilde{G}_{i,j}$ in modelno napovedjo $G(x | \Theta)$. Podobno kot pri prejšnjih CF se razlike najprej kvadrirajo, nato pa povprečijo.

$$of_{surface} = \frac{\sum_{i=1}^{N_{surface}} (\sum_{j=1}^{N_{timestamp}} (\widetilde{G}_{i,j} - G(x | \Theta))^2)}{N_{surface} * N_{timestamp}} \quad (4.6)$$

Omenjena ciljna funkcija ni bila uporabljena v sklopu te naloge.

4.2.2 Združitev ciljnih funkcij in hidravličnega izračuna

Združitev ciljnih funkcij s hidravličnim modelom je idejno enostavna. Najprej idejo opišemo za model s stalnim tokom, kjer je uporabljen samo en časovni interval. Za dan hidravlični model M izračunamo pretoke v ceveh, kjer so bile opravljene meritve $P_{measurmens}$. Temu rečemo modelne napovedi $Flow_{predicted}$. Slednje primerjamo z izmerjenimi pretoki $Flow_{measurmens}$. Razliko kvadriramo, shranimo v tabelo, na koncu pa tabelo povprečimo. Tako dobimo povprečno kvadratno napako pretokov (ang. Mean squared error) MSE_{flows} .

Podoben postopek ponovimo za spoje, kjer so bile opravljene meritve tlakov $J_{measurmens}$. Izmerjene vrednosti $Pressure_{measurmens}$ se primerjajo z modelnimi

napovedmi $Pressure_{predicted}$, razlike se kvadrirajo in povprečijo. Tako dobimo povprečno kvadratno napako pritiskov $MSE_{pressures}$.

Razširitev omenjenega postopka za razširjeno časovno simulacijo je idejno enostavna. Ob prvi iteraciji se hidravlična simulacija izračuna ob času $t = 0$. Funkcija $t = M.run()$ vrne čas trenutnega časovnega koraka. Slednji lahko predstavlja katerega od vmesnih intervalov, kot je opisano v razdelku 2.4.5. Tovrstne časovne korake preskočimo, saj nas zanimajo samo tisti, v katerih so bile opravljene meritve. V vsakem ustreznem koraku nato izračunamo ustrezne razlike med izmerjenimi vrednostmi in modelnimi napovedmi. Enako kot smo storili pri sistemih s stalnim tokom.

Algorithm 4 get differences(model M)

```

 $J_{measurments} = get\_pressure\_measurment\_locations()$ 
 $P_{measurments} = get\_flow\_measurment\_locations()$ 
while true do
     $t = M.run()$  //izracun hidravlicnih razmer ob naslednjem intervalu
    if timestamp_valid(t) then
        for j in  $J_{measurments}$  do
             $Pressure_{measurments} = get\_measured\_pressure(j, t)$ 
             $Pressure_{predicted} = M.get\_pressure(j)$ 
             $\Delta_{pressure}(j, t) = (Pressure_{predicted} - Pressure_{measurments})^2$ 
        end for
        for p in  $P_{measurments}$  do
             $Flow_{measurments} = get\_measured\_flow(p, t)$ 
             $Flow_{predicted} = M.get\_flow(p)$ 
             $\Delta_{flow}(p, t) = (Flow_{predicted} - Flow_{measurments})^2$ 
        end for
    end if
    if t.isLast() then
        Break
    end if
end while

```

4.2.3 Združevanje parametrov

Kot je razvidno iz razdelka 4.1.1, se pri umerjanju srečujemo s problemom razmerja števila neznank in števila meritev. V praksi meritve predstavljajo določeno finančno breme, lahko pa so tudi tehnično zahtevne. Po drugi strani je število cevi v velikih vodovodnih omrežjih lahko zelo veliko. Zato ponavadi ni na voljo dovolj meritev za določitev vseh možnih parametrov cevovoda. S tem se pojavi problem večjega števila neznank, kot je meritev na voljo. Umerjanje takšnega sistema vodi v nedoločljiv sistem enačb oz. pri implicitni metodi umerjanja do možnosti, da se cevem pripišejo nerealne vrednosti.

Po drugi strani lahko imamo natančno enako število meritev, kot je neznanih parametrov, torej cevi v cevovodu. To lahko dosežemo tako, da so meritve opravljane v skoraj vseh vozliščih in povezavah, kar je v praksi nedosegljivo. Po drugi strani lahko na enem merilnem mestu opravimo večje število meritev ob različnih časovnih korakih in tako dosežemo enakost števila neznank in meritev. V praksi se tudi ta pristop izkaže za slabega, saj se zaradi geolokacije merilnih mest in nelinearnosti hidravličnega modela tudi tak sistem izkaže za nedoločljivega.

Zaradi nelinearne zveze med tlaki v vozliščih, hrapavostjo in postavitve merilnih mest, se lahko tudi predoločeni sistemi, kjer imamo na voljo večje število meritev kot parametrov, izkažejo za nedoločljive. Vrednosti parametrov tudi v tem primeru dosegaajo nerealne vrednosti [7].

Da se izognemo velikemu številu meritev, ki so cenovno in tehnično zahtevne, lahko predločen sistem enačb dosežemo na drug način. Z združevanjem parametrov, grupacijo, ustvarimo skupine. Cevi, tako združimo na osnovi nekih lastnosti v skupine z enakimi oziroma vsaj podobnimi karakteristikami. Pri hrapavosti tako združujemo cevi, ki imajo vsaj podobne lastnosti pri:

1. Starosti cevi.
2. Materialu.
3. Geolokaciji v vodovodnem omrežju.

V literaturi tako zasledimo nekaj glavnih načinov grupiranja parametrov, kot je razvidno v tabeli 4.1 [7].

Kriterij grupiranja	Opis kriterija
Britanski kriterij	Cevi grupiramo v skupino, če imajo enake oziroma podobne lastnosti in so topološko povezljive.
Ameriški kriterij	Cevi grupiramo v skupino, če imajo enake oziroma podobne lastnosti.
Britanski senzitivnostni kriterij	Grupiramo samo senzitivne cevovode, ki imajo enake ali podobne lastnosti in so topološko povezljivi.
Ameriški senzitivnostni kriterij	Grupiramo samo senzitivne cevovode, ki imajo enake ali podobne lastnosti.

Tabela 4.1: Načini grupiranja.

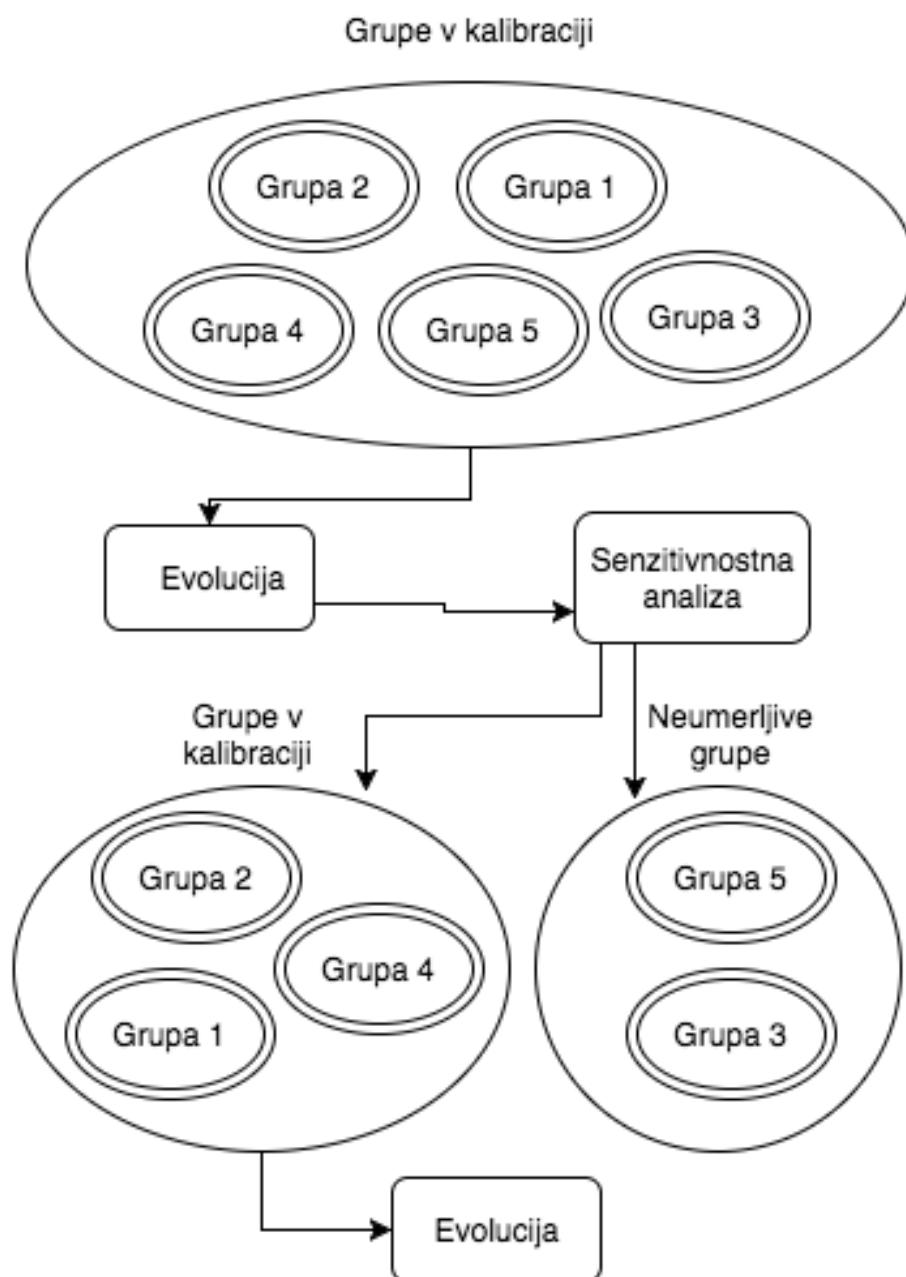
V literaturi sicer pravila o upoštevanju kriterijev za kreiranje grup niso strogo določena, so bolj smernice, končna odločitev pa je prepuščena izdelovalcu hidravličnega modela. Najti je torej potrebno dober kompromis med številom parametrov, njihovo umerljivostjo, kompleksnostjo izračuna, številom meritev, odzivnostjo parametrov in njihovo povezljivostjo [7]. V sklopu te naloge se še posebna pozornost namenjena umerljivosti parametrov.

Kot je bilo nakazano, se srečujemo tudi pri predoločenih primerih umerjanja z nedoločljivostjo parametrov. Posamezna cev oziroma kar celotna skupina je lahko zaradi različnih dejavnikov neumerljiva. Spremembe hrapavosti v tej skupini tako nič oziroma zanemarljivo malo vplivajo na spremembo tlakov na vozliščih in pretokov v ceveh, kjer so bile opravljene meritve. Cevi so lahko bodisi postavljene daleč od merilnih mest, bodisi je njihova dolžina prekratka, da bi imele večji vpliv na model. Aplikacija, ki je bila kreirana v sklopu te naloge, v ta namen omogoča občutljivostno analizo parametrov in z njeno pomočjo izločitev neumerljivih grup iz postopka umerjanja. Grupiranje je še vedno opravljeno na podlagi lastnosti cevi, med postopkom umerjanja pa lahko nekatere grupe izločimo. Slednjim se vrednost hrapavosti določi na podlagi fizikalnih lastnosti cevi. Občutljivostna analiza je podrobneje opisana v nadaljevanju, zaenkrat le predpostavimo, da imamo funkcijo,

ki grupam pripiše njihovo senзитivnost.

Najprej se vse grupe kreirajo v skladu z lastnostmi cevi. Sortirajo se v ločene skupine, kjer imajo vse skupine približno enake lastnosti, kot je razvidno v diagramu 4.2. Postopek se izvede samo enkrat za željen hidravlični model, ki ga umerjamo. Vsaka cev pripade eni skupini, zato se cevem shrani dodaten parameter imenovan skupina. Kalibracijo tako lahko opravimo večkrat z različnimi nastavitvami, pri čemer smo grupacijo opravili le enkrat.

Za določitev senзитivnosti je bil v sklopu te naloge uporabljen povsem nov pristop, ki je opisan v nadaljevanju. Za delovanje potrebuje vsaj nekaj generacij genetskega algoritma, natančneje populacijo, ki je že nekoliko konvergirala proti optimalni rešitvi. Po nekaj iteracijah optimizacije se algoritem ustavi in opravi se senзитivnostna analiza posameznih grup. V tem koraku lahko določimo takoimenovani prag, na podlagi katerega ločimo grupe v dve skupini: tiste, ki so dovolj umerljive za nadaljnje delo in ostale, katerim določimo vrednost hrapavosti in jih ne umerjamo več. Diagram postopka je viden na sliki 4.2.



Slika 4.2: Diagram grupacije.

4.2.4 Povezava parametrov in hidravličnega modela

V prejšnjem razdelku 4.2.3 smo spoznali, da je smiselno združevanje posameznih elementov (cevi) vodovodnega sistema v skupine. Na ta način smo zmanjšali problem in povečali, če ne celo omogočili natančno umerjanje parametrov. Poleg tega smo v razdelku 4.2.2 opisali, kako za dan model M izračunamo vrednosti ciljnih funkcij. V tem razdelku združimo koncept skupin, parametrov Θ in z njihovo pomočjo generiramo model M . Za vsak osebek, ki je predstavljen s parametri Θ , generiramo korespondenčni hidravlični model M .

Parametre Θ zapišemo kot vektor decimalnih števil.

$$\Theta = [P_1 \ P_2 \ \dots \ P_i \ \dots \ P_{N(\Theta)}] \quad (4.7)$$

V vektorju predstavlja i -ti element P_i vrednost hrapavosti cevi v i -ti skupini. Predpriprava modela je idejno enostavna. Opišemo jo z algoritmom 5.

Algorithm 5 prepare model(parameters Θ , Groups)

```

for  $l$  in  $links$  do
  if  $l.isPipe()$  then // ali je povezava cev
    if  $l.inParameters()$  then // ali je cev med povezavami, ki jim
    določamo parametre
      for  $G$  in  $Groups$  do
        if  $l$  in  $G$  then
           $l.roughness() = \Theta_G$ 
        end if
      end for
    end if
  end if
end for

```

V algoritmu pregledamo vse povezave v grafu, ki predstavlja sistem. Najprej izločimo vse povezave, ki niso cevi. V nadaljevanju preverimo, ali je dana cev kandidat za nastavitev hrapavosti. V razdelku 4.2.3 smo omenili, da nekatere skupine in posledično cevi izločimo iz kalibracije in tako zmanjšamo število parametrov in

povečamo umerljivost. Za preostale cevi preverimo, v kateri skupini so vsebovane in jim določimo ustrezno vrednost iz vektorja parametrov Θ .

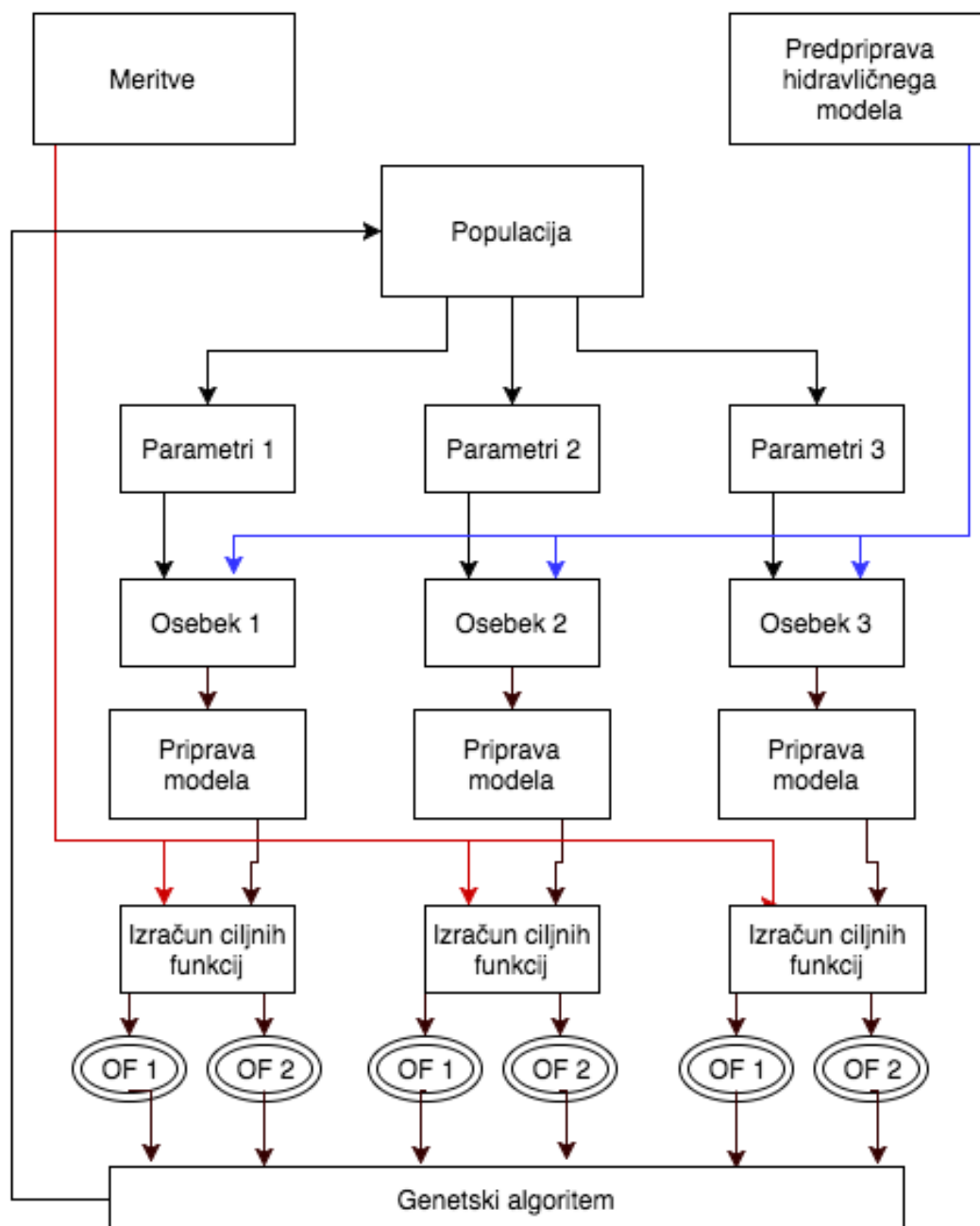
4.2.5 Algoritem umerjanja

Prej omenjene algoritme nad osebki za izračun ciljnih funkcij, opisane v razdelku 4.2.2 in algoritem za pripravo modela iz razdelka 4.2.4 združimo v celoto, ki se izvede v vsakem osebku posebej, kot je razvidno v diagramu 4.3.

Na samem začetku delovanja se najprej predpripravi hidravlični model. Predpriprava je podoben postopek kot algoritem 5, le da se pri tem postopku cevem določijo privzete vrednosti hrapavosti in ne vrednosti iz optimizacije. Na ta način določimo ekspertne vrednosti tudi cevem, ki bodo morda tekom kalibracije izpadle zaradi neumerljivosti. Omenjeni pristop je opisan v razdelku 4.2.3.

Poleg tega se na začetku še uvozijo podatki meritev, ki so shranjene v tekstovnih datotekah. Več o meritvah je napisano v nadaljevanju.

Tekom genetskega algoritma se ob inicializaciji osebkov najprej ustvari ustrezen model na osnovi prejetih parametrov in podatkov grupacije, kot je vidno v algoritmu 5. Nato se izračunajo ciljne funkcije na podlagi popravljenega modela in meritev v algoritmu 4.



Slika 4.3: Diagram umerjanja.

4.3 Analiza kalibracije

4.3.1 Kvantitativni opis kalibracije

Za dan kalibracijski postopek potrebujemo hidravlični model M , meritve v ceveh \tilde{Q} in vozliščih \tilde{h} ter podatke o grupaciji G . Za lažje nadaljnje delo zapišemo nakatere kvalitativne vrednosti, ki bodo opisale hidravlični model.

Spremenljivka	opis
N_{pipes}	število cevi
$N_{junctions}$	število spojev
N_{valves}	število ventilov
N_{tanks}	število vodnih virov
$N_{reservoirs}$	število vodohranov
N_{pumps}	število črpalk

Tabela 4.2: Kvalitativni opis hidravličnega modela.

Vrednosti iz tabele 4.2 vplivajo na časovno zahtevnost hidravličnega izračuna. Večje številke pomenijo počasnejši izračun in na ta način vplivajo na konvergenco. Po drugi strani te vrednosti ne vplivajo veliko na število iteracij genetskega algoritma.

Umerjanje opišemo z vrednostmi, katerih pomen je razviden v tabeli 4.3.

Spremenljivka	opis
$NLOC_{pipes}$	število merilnih mest cevi, kjer se merijo pretoki
$NLOC_{junctions}$	število merilnih mest spojev, kjer se merijo tlaki
$NLOC$	skupno število merilnih mest
$NTIME$	število časovnih korakov
$NMERITEV$	število vseh meritev = $NLOC * NTIME$

Tabela 4.3: Kvalitativni opis meritev.

Večje število meritev poveča časovno zahtevnost, tako pri izračunu hidravličnega modela, kot tudi pri evalvacijskih funkcijah. Po drugi strani večje število meritev vodi v bolj določljiv problem umerjanja, ki praviloma vrača boljše rezultate kalibracije.

Poleg tega lahko dano kalibracijo izvedemo z odstranjevanjem neumerljivih parametrov. Iz tega sledijo vrednosti v tabeli 4.4.

Spremenljivka	opis
N_G	število vseh grup
N_Θ	število grup za umerjanje
$N_{neumerljivi}$	število neumerljivih grup

Tabela 4.4: Kvalitativni opis grupacije.

Vrednosti iz tabele 4.4 najbolj vplivajo na časovno zahtevnost kalibracijskega postopka in njegovo natančnost. Manjše število parametrov N_Θ se odraža v hitrejši konvergenci in bolj določljivemu problemu umerjanja, kar se izraža tudi v večji natančnosti. Če so skupine močno razdrobljene in je N_G večji, se pričakuje neumerljivost pri večjem številu grup.

4.3.2 Izračun Jacobijeve matrike

Jacobijeva matrika je matrika vseh prvih parcialnih odvodov vektorsko definirane funkcije. Potrebujemo jo v nadaljevanju, v tem razdelku pa si pogledimo njen izračun. Najprej zapišimo bolj formalno definicijo.

Izrek 4.1 *Naj bo funkcija f preslikava iz n v m dimenzionalni prostor realnih števil $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, ki za dani vhodni vektor $x \in \mathbb{R}^n$ vrne izhodni vektor $f(x) \in \mathbb{R}^m$. Jacobijeva matrika J za funkcijo f je definirana kot:*

$$J = \frac{\delta f}{\delta x} = \begin{bmatrix} \frac{\delta f}{\delta x_1} & \cdots & \frac{\delta f}{\delta x_n} \end{bmatrix} = \begin{bmatrix} \frac{\delta f_1}{\delta x_1} & \cdots & \frac{\delta f_1}{\delta x_n} \\ \vdots & \ddots & \vdots \\ \frac{\delta f_m}{\delta x_1} & \cdots & \frac{\delta f_m}{\delta x_n} \end{bmatrix} \quad (4.8)$$

Posamezni i, j -ti element lahko zapišemo kot parcialni odvod funkcije f_i po spremenljivki x_j .

$$J_i^j = \frac{\delta f_i}{\delta x_j} \quad (4.9)$$

Jakobijeva matrika je pomembna, saj za dano funkcijo f v okolici točke x opiše linearno aproksimacijo funkcije f v točki x . S pomočjo omenjenih aproksimacij v nadaljevanju razvijemo več evalvacijskih procesov. Brez večjih matematičnih izpeljav si pogledajmo izračun Jacobijeve matrike za primer umerjanja parametrov vodovodnih sistemov.

Kot je bilo predhodno opisano, lahko za dan hidravlični model M , nabor parametrov θ , množico izmerjenih pretokov \tilde{Q} in množico izmerjenih tlakov \tilde{H} izračunamo vrednost razlike pretokov $f_Q(\theta)$ ter vrednost razlike tlakov $f_H(\theta)$. Kot je vidno v algoritmu 4, lahko za vsako merilno mesto tlaka j ali pretoka p ob danem časovnem koraku t izračunamo vrednost modelne napovedi, $H(j, t)$ in $Q(p, t)$. Omenjeni matriki, katerih velikosti sta $(j \times t)$ in $(p \times t)$, najprej prestrukturiramo v vektor velikosti $(j * t + p * t \times 1)$, katerega sestava je opisana v formuli 4.10.

$$D(\Theta) = \begin{bmatrix} V_{pressure} \\ V_{flow} \end{bmatrix} \quad (4.10)$$

Vektor sestavlja $V_{pressure}$ in V_{flow} , ki sta sama po sebi tudi vektorja, sestavljena, kot je vidno v 4.11 in 4.12.

$$V_{pressure} = \begin{bmatrix} v_{pressure,1} & v_{pressure,2} & \dots & v_{pressure,i} & \dots & v_{pressure,t} \end{bmatrix}^T \quad (4.11)$$

$$V_{flow} = \begin{bmatrix} v_{flow,1} & v_{flow,2} & \dots & v_{flow,i} & \dots & v_{flow,t} \end{bmatrix}^T \quad (4.12)$$

Ob i -tem časovnem intervalu posamezen člen v $V_{pressure}$ predstavlja $v_{pressure,i}$ opisani v formuli 4.13. Člen ob i -tem časovnem intervalu vektorja V_{flow} je $v_{flow,i}$, opisani v 4.14.

$$v_{pressure,i} = \begin{bmatrix} H(1,i) & H(2,i) & \dots & H(j,i) \end{bmatrix} \quad (4.13)$$

$$v_{flow,i} = \begin{bmatrix} Q(1,i) & Q(2,i) & & Q(p,i) \end{bmatrix} \quad (4.14)$$

Na ta način smo za dan model M pri parametrih Θ pridobili vektor $D(\Theta)$, ki bo osnova za računanje Jacobijeve matrike. Sedaj lahko podobne vektorje $D(\tilde{\Theta}_i)$ izračunamo tudi za vektorje parametrov $\tilde{\Theta}_i$.

Vektor parametrov $\tilde{\Theta}_i$ dobimo iz prvotnega vektorja parametrov Θ tako, da i -ti element pomnožimo z 1.1, vse ostale pa pustimo takšne kot so. Primer takšnega vektorja je viden v tabeli 4.5.

	P_1	P_2	P_3	P_4	P_5	P_6
Θ	0.5	2.2	2.0	6.0	8.0	2.2
$\tilde{\Theta}_2$	0.5	2.42	2.0	6.0	8.0	2.2
$\tilde{\Theta}_3$	0.5	2.42	2.2	6.0	8.0	2.2

Tabela 4.5: Primer vektorjev $\tilde{\Theta}_i$.

Sedaj lahko iz originalnega vektorja parametrov Θ in korespondenčnega vektorja $D(\Theta)$ ter vektorjev s popravljenim parametrom $\tilde{\Theta}_i$ in korespondenčnih $D(\tilde{\Theta}_i)$ končno izračunamo Jacobijevo matriko.

Posamezni element matrike predstavlja odvod razlike Δ_{flow} oziroma $\Delta_{pressure}$ na danem merilnem mestu ob danem časovnem koraku, po enem izmed parametrov vektorja $D(\Theta)$. Izračun odvoda je možen na več načinov, za ta namen pa je uporabljena enostavna dvotočkovna aproksimacija, ki izračuna naklon premice skozi točki $(x, f(x))$ in $(x + h, f(x + h))$. Pri dovolj majhni vrednosti h dobimo dovolj natančno vrednost odvoda.

$$\frac{f(x + h) - f(x)}{h} \quad (4.15)$$

Odvod pri j -tem elementu vektorja $D(\Theta)$ po i -tem parametru vektorja Θ izračunamo kot:

$$\frac{\|D(\Theta, j) - D(\widetilde{\Theta}_i, j)\|}{\|\Theta(i) - \widetilde{\Theta}_i(i)\|} \quad (4.16)$$

V formuli sta $D(\Theta, j)$ in $D(\widetilde{\Theta}_i, j)$ j -ta elementa vektorjev $D(\Theta)$ in $D(\widetilde{\Theta}_i)$. $\Theta(i)$ je prvotna vrednost i -tega elementa vektorja parametrov Θ in $\widetilde{\Theta}_i(i)$ korendenčna popravljena vrednost.

Na ta način dobimo Jacobijevo matriko. Število stolpcev sovpada s številom parametrov N_Θ , število vrstic pa s skupnim številom meritev $j*t + p*t$. Omenjeno matriko lahko izračunamo za vsako možno rešitev, ker pa gre za časovno zahtevno operacijo, jo izračunamo samo za izbrano rešitev končne populacije. Gre za časovno najzahtevnejši algoritem v tej nalogi. Hidravlično simulacijo izračunamo N_Θ krat, pri čemer za posamezno simulacijo potrebujemo $\mathcal{O}(n^{2.807})$. Skupna časovna zahtevnost je torej $\mathcal{O}(N_\Theta * n^{2.807})$.

4.3.3 Stopnja ujemanja modela

Prva statistika, ki jo lahko povemo o umerjanju hidravličnega modela, je stopnja ujemanja modela. Zanima nas, v kolikšni meri se meritve ujemajo z modelnimi napovedmi. V literaturi se pojavlja kup različnih regresijskih statistik, ki opisujejo ujemanje. Eno od njih smo že spoznali in je uporabljena v samem kalibracijskem postopku - MSE (ang. mean square error). Uporabljena je v obeh ciljnih funkcijah ter vrača povprečno kvadratno odstopanje razlike med izmerjenimi in modelnimi pretoki in tlaki.

$$of_{pressure} = \frac{\sum_{i=1}^{N_{pressure}} (\sum_{j=1}^{N_{timestamp}} (\widetilde{H}_{i,j} - H(x | \Theta))^2)}{N_{pressure} * N_{timestamp}} \quad (4.17)$$

$$of_{flow} = \frac{\sum_{i=1}^{N_{flow}} (\sum_{j=1}^{N_{timestamp}} (\widetilde{Q}_{i,j} - Q(x | \Theta))^2)}{N_{flow} * N_{timestamp}} \quad (4.18)$$

Za verifikacijo ujemanja ocenimo povprečno kvadratno napako (MSE) tudi na množici verifikacijskih vozlišč in cevi. Pri tem načinu model kalibramo s

pomočjo kalibracijske množice meritev, po koncu umerjanja pa preverimo ujemanje s pomočjo dodatne množice neodvisnih meritev. Izračun verifikacijskih povprečnih kvadratnih napak je identičen izračunu ciljnih funkcij. Verifikacijsko MSE pretokov v nadaljevanju označimo z $MSE_{flow}^{validation}$, verifikacijsko MSE pritiskov pa $MSE_{pressure}^{validation}$.

Bolj napredne statistike upoštevajo tudi razmerje števila parametrov in meritev. V nalogi je uporabljena varianca regresijske napake.

$$s^2 = \frac{cf(\Theta)}{v} \quad (4.19)$$

V formuli je $f(\Theta)$ vrednost poljubne ciljne funkcije ali pa vrednost skupne povprečne ciljne funkcije. Če uporabimo ciljno funkcijo pretokov cf_{flow} dobimo varianco regresijske napake s_{flow}^2 , pri ciljni funkciji tlakov $cf_{pressure}$ pa $s_{pressure}^2$. Če uporabimo eno samo ciljno funkcijo, je notacija za varianco regresijske napake kar s^2 . Vrednost v predstavlja število prostorskih stopenj.

$$v = NLOC * NTIME - N_{\Theta} = NMERITEV - N_{\Theta} \quad (4.20)$$

V enačbi je $NLOC$ število merilnih mest, $NTIME$ število časovnih korakov in N_{Θ} število parametrov umerjanja.

4.3.4 Ocena negotovosti s FOSM

Drugo pomembno vprašanje, na katerega moramo odgovoriti pri umerjanju, je ocena negotovosti ujemanja parametrov in njihova odzivnost. Ker je optimizacijski proces z genetskimi algoritmi bolj ali manj naključen proces, ki deluje brez kakršnegakoli predznanja o vrednostih parametrov ali vsaj njihovih gradientih, se negotovost in z njo povezane vrednosti parametrov odražajo v samih rezultatih optimizacije. Opazimo, da nekateri parametri in skupine, ki jih slednji opisujejo, nosijo veliko informacijsko težo. Z drugimi besedami, majhna odstopanja parametrov od realnih vrednosti hrapavosti se odražajo v velikih odstopanjih med meritvami in modelnimi napovedmi, kar ima velik vpliv na ciljne funkcije. Po drugi strani nekateri parametri in skupine skorajda nimajo vpliva na vrednost

ciljne funkcije. Hrapavosti teh skupin lahko zasedejo širše območje vrednosti, brez večjega vpliva na ciljno funkcijo.

V tem razdelku brez dodatne matematične izpeljave opišemo metodo za oceno odzivnosti FOSM (ang. first order second moment). FOSM je metoda za oceno odzivov in variance funkcije, ki temelji na uporabi odvoda in Taylorjeve vrste. Metoda predpostavlja, da so parametri slučajne spremenljivke z normalno porazdelitvijo. Prek kovariančne matrike tako opišejo porazdelitev parametra. Ker se pri izpeljavi uporabijo le prvi členi Taylorjeve vrste, se predvideva linearna zveza med parametri in ciljnim funkcijami, kar predstavlja določeno omejitev [7].

Jedro te metode je variančno-kovariančna matrika Σ , ki jo izračunamo s pomočjo Jacobijeve matrike, opisane v razdelku 4.3.2. Variančno kovariančna matrika je kvadratna in vsebuje variance in kovariance povezane s parametri. Diagonalni elementi matrike vsebujejo variance parametrov, nediagonalni elementi pa kovariance med vsemi možnimi pari parametrov. Matrika je simetrična, saj je kovarianca med parametroma A in B enaka kovarianci med B in A [21]. Velja pravilo $Var(x) = Cov(x, x)$ [20]. Njeno aproksimacijo izračunamo s pomočjo formule 4.21.

$$\Sigma = S^2 * (J^T * W * J)^{-1} \quad (4.21)$$

V enačbi je s^2 prej omenjena varianca regresijske napake, W kvadratna matrika utežnih koeficientov. V sklopu te naloge utežni koeficienti niso bili uporabljeni, zato je W kar matrika identiteta. J in J^T sta Jacobijevi matriki, opisani v sekciji 4.3.2 in zavzemata vlogo senzitivnostne matrike. Podaja približek odzivnosti razlike meritev in modelnih napovedi v odvisnosti od spremembe parametra.

Diagonalni elementi variančno-kovariančne matrike Σ tako predstavljajo variance istoležnih parametrov.

$$\sigma_{\Theta,k}^2 = \Sigma_{k,k} \quad (4.22)$$

S pomočjo variance posameznih parametrov lahko izračunamo interval zaupanja.

$$\Theta_k - t(n, 1 - \alpha/2) * \sigma_{\Theta,k}^2 \leq \Theta_k \leq \Theta_k + t(n, 1 - \alpha/2) * \sigma_{\Theta,k}^2 \quad (4.23)$$

V formuli je $t(n, 1 - \alpha/2)$ studentova porazdelitev in α koeficient tveganja.

Z metodo FOSM tako prek variance pridemo do intervalov zaupanja. Večja varianca pomeni širši interval zaupanja in posledično nižjo umerljivost parametra. Na isti način ocenimo tudi odzivnost parametra. Parametri z večjo informacijsko vrednostjo so bolj umerljivi, imajo večji vpliv in bodo bolj konvergirali proti globalnemu optimumu, kar se bo prek večjega parcialnega odvoda v Jacobijevi matriki odražalo v nižji varianci parametra.

Metoda FOSM ima visoko časovno zahtevnost, saj za delovanje potrebuje izračun Jacobijeve matrike s časovno zahtevnostjo $\mathcal{O}(N_{\Theta} * n^{2.807})$, kjer je n število cevovodov oziroma vozlišč vodovodnega sistema in N_{Θ} število parametrov. Zahtevnost tako narašča s kubično velikostjo sistema. V nadaljevanju potrebujemo še računanje inverza matrike $N_{\Theta} \times N_{\Theta}$, katere zahtevnost je $\mathcal{O}(N_{\Theta}^{2.807})$. Poleg tega potrebujemo še množenje matrik z zahtevnostjo $\mathcal{O}(N_{\Theta} * N_{\Theta} * N_{MERITEV})$. Skupna časovna zahtevnost je enaka zahtevnosti izračuna Jacobijeve, torej $\mathcal{O}(N_{\Theta}^{2.807})$.

Da se izognemo računanju inverza matrike in tako pridobimo na času in numerični stabilnosti, lahko do varianc parametrov pridemo samo z množenjem vektorjev.

$$\sigma_{\Theta,k}^2 = s^2 * \frac{1}{J_k \cdot J_k} \quad (4.24)$$

V formuli je J_k k -ti stolpec Jacobijeve matrike in s^2 varianca regresijske napake. Uporabljena je bila formula za izračun inverza diagonalne matrike.

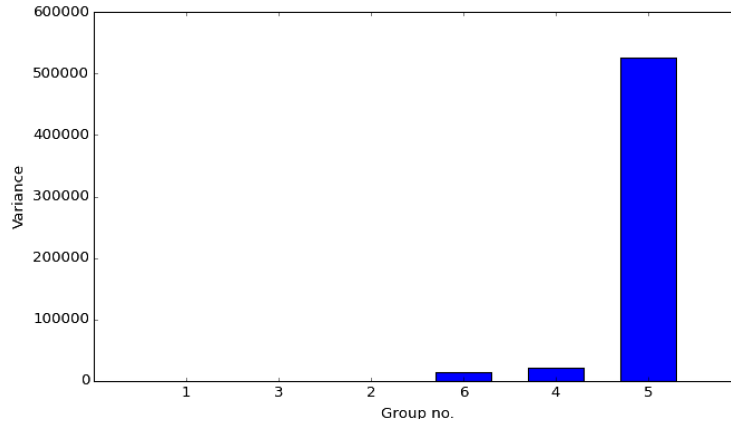
$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \frac{1}{a_{nn}} \end{bmatrix} \quad (4.25)$$

Metoda FOSM potrebuje za izračun dobre senzitivnostne (Jacobijeve) matrike, dobro umerjen model. To je tudi največja pomankljivost metode. V razdelku o grupaciji parametrov 4.2.3 smo povedali, da potrebujemo metodo za analizo občutljivosti parametrov na začetku kalibracije, da odstranimo parametre, ki niso umerljivi. Ker metoda FOSM potrebuje za delovanje že umerjen hidravlični model, v ta namen ne bo najprimernejša. Na sliki 4.4 so vidne variance parametrov na

neumerjenem modelu. Vidno je, da se variance precej razlikujejo od primera z umerjenim modelom na sliki 4.5.

Izračun je lahko v nekaterih primerih tudi numerično nestabilen. Pri zelo nizki informacijski vrednosti parametra so odvodi razlik meritev in modelnih napovedi po parametru blizu ničle. Tudi skalarni produkt vektorjev takšnih odvodov $J_k \cdot J_k$ je posledično enak nič. Na ta način hitro preidemo v deljenje z zelo majhnim številom, kar se odraža v visokih vrednostih variance.

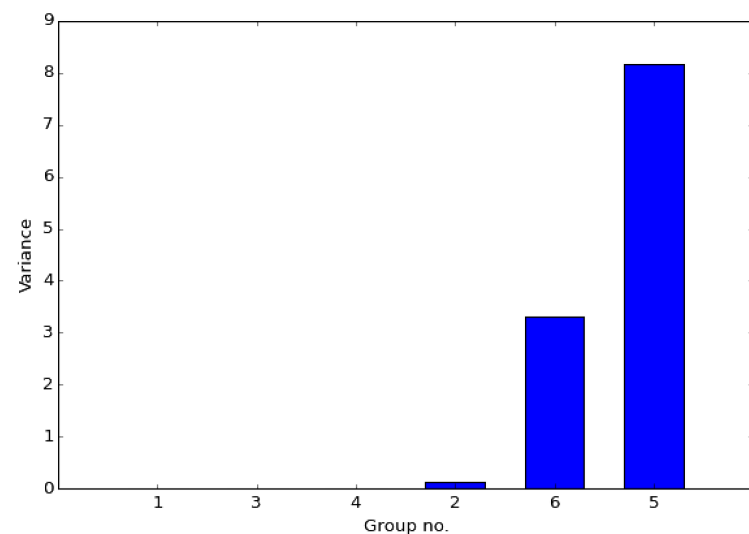
Metoda predpostavlja, da so odstopanja meritev od modelnih napovedi normalno porazdeljena, kar pa v realnih primerih ni nujno res. Vrednosti variančno-kovariančne matrike so poleg tega odvisne tudi od vrednosti variance regresijske napake, ki je lahko v primeru nizkih vrednosti ciljne funkcije in velikega števila prostorskih stopenj zelo nizka.



Slika 4.4: Variance po posameznih grupah z metodo FOSM na neumerjenem modelu.

Na podlagi izračunanih vrednosti varianc lahko definiramo statistiko, ki ocenjuje skupno negotovost parametrov.

$$\hat{\Sigma} = Trace(\Sigma) = \sum_{k=1}^{N_{\Theta}} \sigma_{\Theta,k}^2 \quad (4.26)$$



Slika 4.5: Variance po posameznih grupah z metodo FOSM na umerjenem modelu.

4.3.5 Ocena odzivnosti in negotovosti z GA

V razdelku 4.2.3 smo omenili, da potrebujemo za vsak parameter Θ_k podatek o njegovi umerljivosti. Na ta način lahko izločimo neumerljive parametre in izboljšamo optimizacijski proces. V prejšnjem razdelku smo utemeljili, da metoda FOSM v ta namen ni najboljša, saj potrebuje že dobro umerjen model za delovanje.

Med optimizacijskim postopkom lahko opazimo, da nekateri parametri hitreje zavzamejo vrednost okoli optimalne, drugi pa potrebujejo večje število generacij za konvergenco. Ideja ocenjevanja negotovosti s pomočjo genetskih algoritmov temelji na točno tem opazanju. S statistično analizo populacije ugotovimo raztros vrednosti po posameznem parametru. Orodje v sklopu te naloge shranjuje vrednosti genomov v matriki.

$$STAT_{zadnjapopulacija} = \begin{bmatrix} \Theta(1) \\ \Theta(2) \\ \vdots \\ \Theta(np) \end{bmatrix} = \begin{bmatrix} \Theta(1)_1 & \Theta(1)_2 & \dots & \Theta(1)_{N_\Theta} \\ \Theta(2)_1 & \Theta(2)_2 & \dots & \Theta(2)_{N_\Theta} \\ \vdots & \vdots & & \vdots \\ \Theta(NP)_1 & \Theta(NP)_1 & \dots & \Theta(NP)_{N_\Theta} \end{bmatrix} \quad (4.27)$$

V matriki je v i -ti vrstici gen $\Theta(i)$, ki ga sestavljajo parametri $\Theta(i)_1, \Theta(i)_2, \dots, \Theta(i)_{N_\Theta}$. Število vrstic v matriki je enako velikosti populacije, število stolpcev pa številu parametrov. Podobno matriko se vodi za vse gene, ki so bili kadarkoli uporabljeni $STAT_{vsepopulacije}$.

Iz omenjenih matrik se nato izračuna varianca po posameznem stolpcu iz formule za standardni odklon.

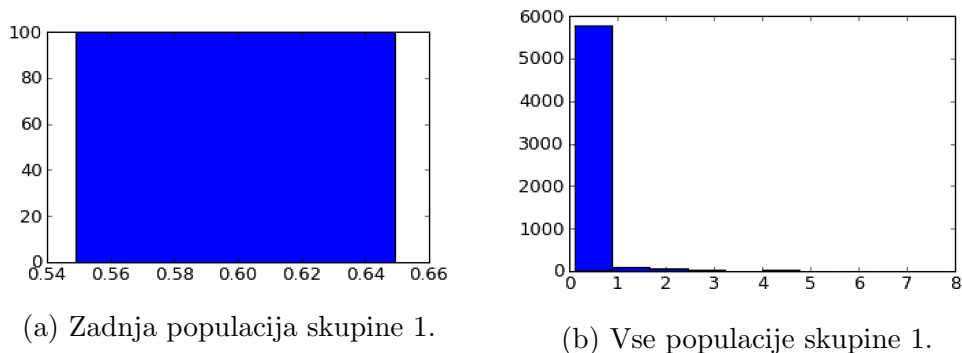
$$\sigma_k^2 = \frac{1}{NP} \sum_{i=1}^{NP} (\Theta(i)_k - \mu_k)^2 \quad (4.28)$$

V formuli je σ_k^2 varianca k -tega parametra in μ_k povprečna vrednost k -tega parametra.

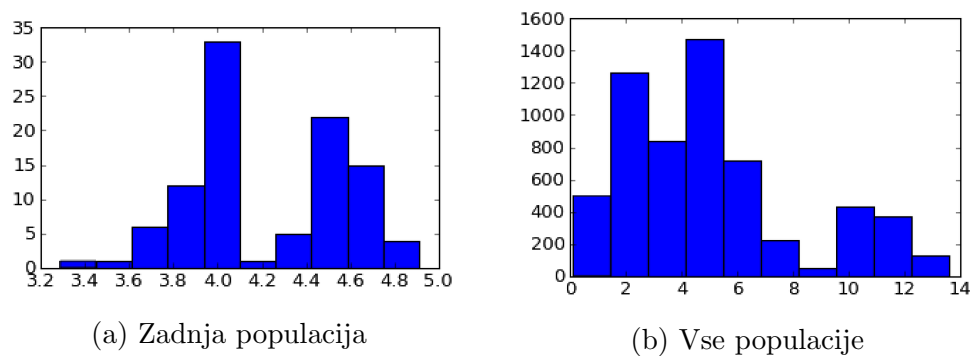
Prednost te metode je, da ne potrebujemo dodatnih kalkulacij, uporabimo samo podatke, ki se generirajo tekom delovanja genetskega algoritma. Po poteku nekaj generacij algoritem ustavimo in prikažemo varianco vseh populacij in zadnje. Pri parametrih, kjer je prišlo do konvergence se pokaže vidna razlika, kot je vidno na sliki 4.8.

Slaba stran te metode je njena odvisnost od inicializacije populacije. Ker se vrednosti parametrov na začetku generirajo naključno lahko tako vplivajo na hitrost konvergence različnih parametrov.

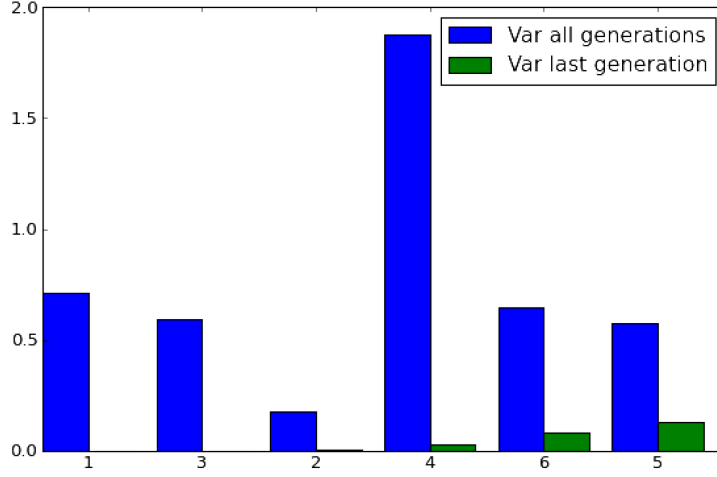
Za dodatne informacije, lahko prikažemo tudi histogram porazdelitve posameznega parametra, kot je vidno na slikah 4.6 in 4.7.



Slika 4.6: Histograma grupe 1 zadnje populacije (a) in vseh populacij (b). Iz (b) je razvidno, da so vrednosti hitro konvergirale. V zadnji populaciji (a) so vrednosti zbrane v ozkem območju okoli 0.6.



Slika 4.7: Histograma grupe 6 zadnje populacije (a) in vseh populacij (b). Iz (b) je razvidno, da so vrednosti tekom optimizacijskega postopka zasedale skoraj celotno definicijsko območje. V zadnji populaciji (a) so vrednosti zbrane v širšem območju od 3.2 do 5.0.



Slika 4.8: Variance po posameznih grupah z analizo populacije na umerjenem modelu.

4.3.6 Ocena pristranskosti

Po optimizaciji pričakujemo, da bodo odstopanja med meritvami in modelnimi čim manjša, kar se bo kazalo v visoki stopnji ujemanja modela iz razdelka 4.3.3. Pri analizi ujemanja nas poleg stopnje ujemanja zanima še porazdelitev. V tem razdelku obravnavamo statistiko razlik modelnih napovedi in meritev.

Prvi statistični kazalnik, ki opisuje odstopanja, je bias. Pri dovolj dobrem ujemanju pričakujemo, da bo povprečje odstopanj približno 0. Postopek izračuna biasa je podoben izračunu ciljne funkcije, le da se odstopanje ne kvadrira.

$$BIAS_{pressure} = \frac{\sum_{i=1}^{N_{pressure}} (\sum_{j=1}^{N_{timestamp}} (\widetilde{H}_{i,j} - H(x | \Theta)))}{N_{pressure} * N_{timestamp}} \quad (4.29)$$

$$BIAS_{flow} = \frac{\sum_{i=1}^{N_{flow}} (\sum_{j=1}^{N_{timestamp}} (\widetilde{Q}_{i,j} - Q(x | \Theta)))}{N_{flow} * N_{timestamp}} \quad (4.30)$$

Poleg povprečja odstopanj nas zanima še nekaj drugih lastnosti odstopanj tlakov $\Delta_{pressure}$ in pretokov Δ_{flow} . V tabeli 4.6 so našteje še ostale statistične ocene odstopanj.

Statistična ocena	Opis
$MAX(\Delta_{flow})$	Maksimalno odstopanje pretokov
$MIN(\Delta_{flow})$	Minimalno odstopanje pretokov
$MAX(\Delta_{pressure})$	Maksimalno odstopanje tlakov
$MIN(\Delta_{pressure})$	Minimalno odstopanje tlakov
$\sigma(\Delta_{flow})$	Standardni odklon odstopanj pretokov
$\sigma(\Delta_{pressure})$	Standardni odklon odstopanj tlakov

Tabela 4.6: Enostavnejše statistične ocene odstopanj modelnih napovedi od meritev.

Poleg naštetih ocen je v sklopu naloge uporabljena še ocena relativnega biasa, ki podaja relativno natančnost ujemanja. V idealnih pogojih se podobno kot pri biasu pričakuje vrednost 0, vendar se v praksi pričakujejo manjša odstopanja.

$$RELBIAS = \frac{BIAS}{\sigma(\Delta)^2} \quad (4.31)$$

Poglavje 5

Umerjanje na primerih

V sklopu tega poglavja je predstavljeno umerjanje parametrov vodovodnega sistema na več primerih vodovodnega omrežja.

5.1 Priprava meritev

Umerjanje v sklopu te naloge je opravljeno na dveh različnih hidravličnih modelih, manjšem Anytown AM in večjem Aquas de Coimbra. Zaradi poenostavitve dela in lažje verifikacije in testiranja so bile v tej nalogi uporabljene statične meritve. Za dana dva, dobro delujoča in pregledana modela smo generirali meritve. V nalogi je bil v ta namen sestavljen dodaten program za zajem meritev, katerega delovanje je idejno enostavno. Uporabnik izbere željeno merilno mesto, bodisi vozlišče za tlak, bodisi cev za pretok na delujočem hidravličnem modelu. Program nato zažene hidravlično simulacijo. Na vsakem časovnem koraku se preveri ustreznost koraka podobno, kot je opisano v algoritmu 4. Pri vsakem ustreznem časovnem koraku se nato odčita modelna napoved željene hidravlične veličine. Da meritve bolj posnemajo situacijo v naravi, se je meritvam dodal še šum z normalno porazdelitvijo.

$$f(x \mid \sigma^2, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (5.1)$$

Tu je bil uporabljen raztros $\sigma^2 = 0.5$ in $\mu = 0$. Na ta način smo posnemali napake pri merjenju v naravi, do katerih vedno prihaja. Poleg tega se s tem preverja robustnost algoritma.

Zaradi uporabne predpripravljenih modelov je bilo delo na testnih primerih enostavnejše.

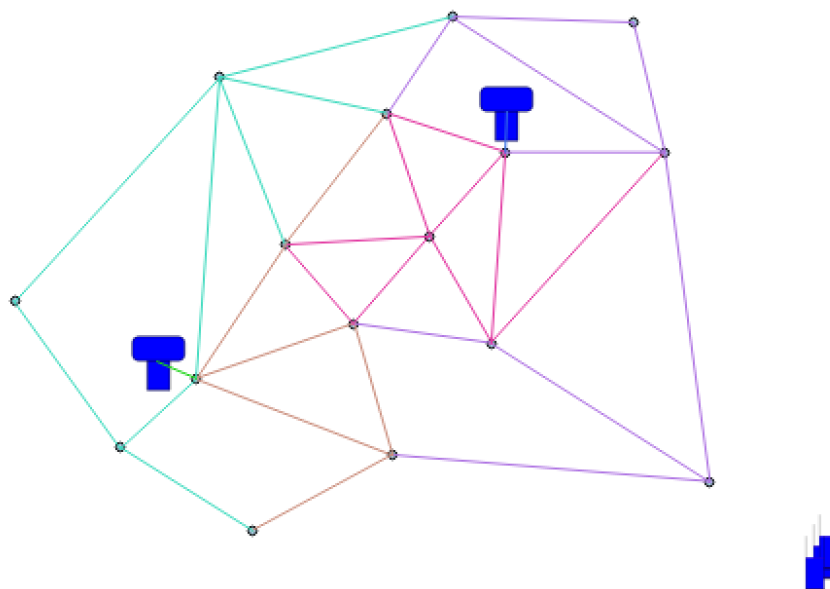
5.2 Uporabljeni modeli

Kot je bilo omenjeno, sta bila v nalogi uporabljena dva hidravlična modela, izhajajoča iz dveh zelo različnih vodovodnih omrežij. Prvi je manjši hidravlični model Anytown AM. Gre za model izmišljenega vodovodnega sistema, ki je pogosto uporabljen v literaturi. Namenjen je testiranju različnih algoritmov. Tudi v tej nalogi je bila večina testiranja opravljena na omenjenem modelu. Zaradi relativno majhnega števila povezav in uporabljenih časovnih korakov je hidravlični izračun dokaj hiter in tako primeren za testiranje. Kvantitativni opis vodovodnega omrežja je razviden v tabeli 5.1.

Spremenljivka	Število elementov
N_{pipes}	34
$N_{junctions}$	16
N_{valves}	0
N_{tanks}	2
$N_{reservoirs}$	3
N_{pumps}	3

Tabela 5.1: Kvantitativni opis hidravličnega modela Anytown AM.

Drug uporabljen model je precej kompleksnejši. Gre za hidravlični model dela Portugalskega mesta Coimbra. Podatke omrežja je pripravilo podjetje Aquas de Coimbra, ki upravlja in vzdržuje tamkajšnji vodovodni sistem [22] in so prosto dostopni v sklopu QGIS-ovskega vtičnika Oslandia. Orodje QGIS, v sklopu katerega je implementirano orodje za kalibracijo, je opisano v nadaljevanju. Slednje omogoča vzpostavitev EPANET-ovih .inp datotek (razdelek 2.5), iz prej omenjenih podatkov geografsko-informacijskega sistema. Znotraj geografsko informacijskega sistema QGIS je omenjeno vodovodno omrežje lepo vidno kot del istoimenskega portugalskega mesta, kot je vidno na sliki 5.2.



Slika 5.1: Shema vodovodnega omrežja Anytown AM.

Iz slike je razvidno, da gre za obsežen vodovodni sistem, ki je kvantitativno opisan v tabeli 5.2. Sistem obsega 29.27 km cevi. Hidravlični model, katerega podatke je pripravilo isto podjetje, opisuje hidravlično dogajanje enega dne. Uporabljen je korak razširjene časovne simulacije 5 minut.

Spremenljivka	Število elementov
N_{pipes}	294
$N_{junctions}$	272
N_{valves}	1
N_{tanks}	1
$N_{reservoirs}$	1
N_{pumps}	1

Tabela 5.2: Kvalitativni opis hidravličnega modela Aquas de Coimbra.

5.3 Primer: Anytown AM

Prvi primer umerjanja je prikazan na enostavnejšem modelu Anytown AM. Grupacija je izvedena tako, da so cevi znotraj grupe povezane, kar je precej enostavno zaradi močno povezanega omrežja. Tabela 5.3 prikazuje cevi, ki sestavljajo posamezno od šestih nastalih grup. Grupacija je lepo vidna tudi na sliki 5.1, kjer so cevi obarvane v skladu z grupo.

Grupa	Elementi grupe
1	P2, P6, P4, P10, P30, P32, P36, P34, P38
3	P52, P50, P48, P46, P28, P44,
4	P56, P58, P42, P40, P66, P62, P60, P64,
2	P8, P12, P24, P26, P22, P16, P14, P18, P20
5	P80
6	P78

Tabela 5.3: Grupacija.

Meritve za primer Anytown AM so zajete na več merilnih mestih, kot je vidno v tabeli 5.4. Delež merilnih mest pretokov je 5.8%, tlakov pa 25%. Časovni potek pretokov na merilnih mestih je viden na grafu 5.3, tlakov pa na grafu 5.4.

Tip merilnega mesta	Število	Merilna mesta
Meritev pretoka	2	P78, P80
Meritev tlaka	4	J40, J90, J120, J140
Verifikacijska meritev pretoka	2	P28, P4
Verifikacijska meritev tlaka	2	J20, J170

Tabela 5.4: Merilna mesta v modelu Anytown AM.

Pri genetskem algoritmu je uporabljena velikost populacije 100 osebkov, kjer se v vsaki generaciji generira 50 otrok. Uporabljena verjetnost mutacije je 8%. Območje preiskovanja, znotraj katerega lahko parametri zavzamejo vrednost, je od

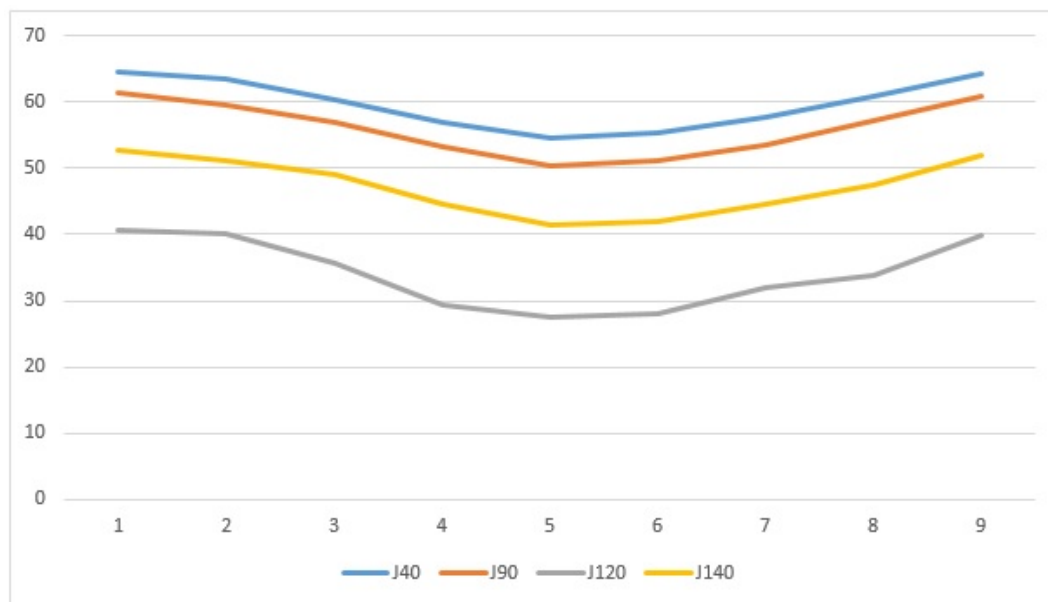


Slika 5.3: Merilna mesta pretokov in njihovi podatki v različnih časovnih korakih.

0.01 do 20.0. Število otrok neposredno vpliva na hitrost evalvacije ene generacije. Več otrok pomeni več evalvacij in počasnejše izvajanje, po drugi strani pa večjo možnost, da najdemo ugodno rešitev. Med testiranjem se je iz uporabniške izkušnje izkazalo to število otrok kot ugodno. Manjše pomeni več generacij izvajanja, večje pa počasnejšo evalvacijo ene generacije. Razmerje med velikostjo populacije in številom otrok tudi vpliva na zaustavitveni kriterij. Če je število otrok skoraj enako veliko kot velikost populacije je sprememba med generacijami večja, torej bo zaustavitveni kriterij dosežen kasneje.

Med delovanjem GA se ciljni funkciji najprej hitro približata optimumu, nato pa se konvergenca upočasni, kot je vidno na sliki 5.5. Uporabljen je ustavitveni pogoj $\varepsilon = 0.001$ tako za pretoke kot za tlake, pri čemer se optimizacija ustavi v 90 generacijah.

Rezultati optimizacije so zaradi enostavnosti omrežja in dobrega razmerja med velikostjo in številom merilnih mest dobri, kot je vidno v tabeli 5.5. Stopnja ujemanja modela je visoka. Da umerjeni model res ponazarja realno stanje, se

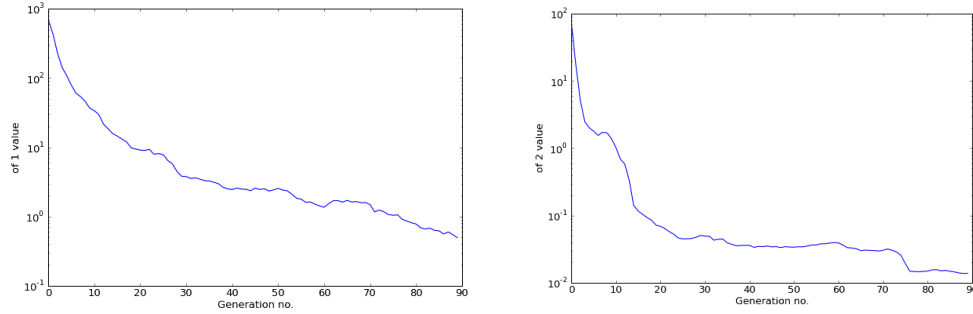


Slika 5.4: Merilna mesta tlakov in njihovi podatki v različnih časovnih korakih.

kaže tudi prek nizkih vrednosti validacijskih ciljnih funkcij. Ujemanje je torej splošno in ne samo na izbranih merilnih mestih.

Parameter skupne negotovosti parametrov $\hat{\Sigma}$ nakazuje, da ni nobenega večjega nezaupanja v vrednost kateregakoli parametra. Analiza odstopanj kaže manjšo pristranskost, širina raztrosa pa je brez večjih posebnosti.

V tabeli 5.6 so vidni rezultati optimizacije, torej vrednosti parametrov hrapavosti za posamezno grupo. Poleg njih so navedene tudi variance po metodi FOSM in korespondenčni intervali zaupanja. Ker uporabljamo delujoč model za kalibracijo, so za preverjanje delovanja prikazane tudi povprečne vrednosti hrapavosti dejanskega modela.



(a) Povprečna vrednost of_{flow} po generacijah. (b) Povprečna vrednost $of_{pressure}$ po generacijah.

Slika 5.5: Vrednosti ciljnih funkcij po generacijah.

Parameter	Vrednost
of_{flow}	0.0039
$of_{pressure}$	0.0001
$MSE_{flow}^{validation}$	0.0099
$MSE_{pressure}^{validation}$	$3.30 * 10^{-5}$
s^2	0.0016
s_{flow}^2	$8.29 * 10^{-5}$
$s_{pressure}^2$	$3.07 * 10^{-6}$
$\hat{\Sigma}$	0.0628
$BIAS_{flow}$	-0.0046
$BIAS_{pressure}$	0.0011
$MAX(\Delta_{flow})$	0.1239
$MIN(\Delta_{flow})$	-0.1204
$MAX(\Delta_{pressure})$	0.0297
$MIN(\Delta_{pressure})$	-0.0228
$\sigma(\Delta_{flow})$	0.0629
$\sigma(\Delta_{pressure})$	0.0121
$RELBIAS_{flow}$	-0.0740
$RELBIAS_{pressure}$	0.0963

Tabela 5.5: Analiza umerjanja modela Anytown AM.

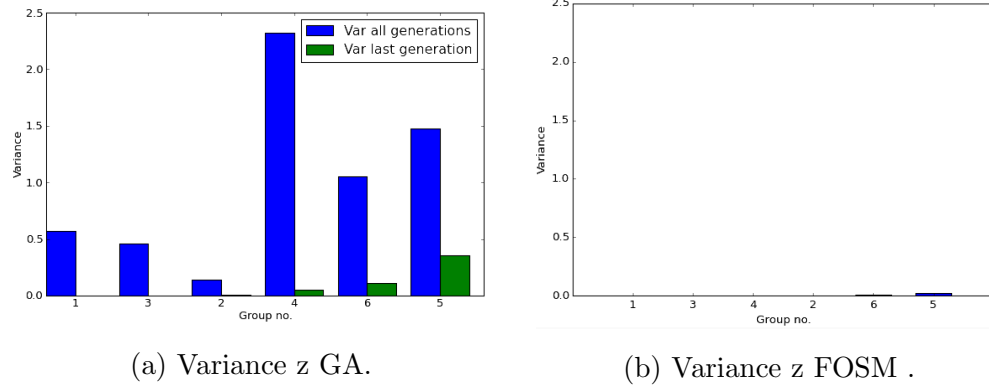
Grupa	Θ_i	Povprečna hrapavost	Varianca	Interval od	Interval do
1	0.5172	0.525	$3.29 * 10^{-7}$	0.5172	0.5172
3	2.5589	2.5	$7.84 * 10^{-6}$	2.5589	2.5590
4	0.2992	0.3	$2.24 * 10^{-5}$	0.2992	0.2993
2	11.9765	11.75	0.00095	11.9747	11.9782
6	0.5419	1.2	0.01280	0.5168	0.5670
5	0.8010	1.2	0.0491	0.7047	0.8973

Tabela 5.6: Rezultati umerjanja modela Anytown AM.

Zanimiva je tudi primerjava metod določanja umerljivosti, torej metode na osnovi FOSM in metode s statistično analizo populacije. Na grafu 5.6 so vidne variance po grupah z metodo FOSM in variance parametrov populacije. Jasno je vidno, da obe metodi kažeta kot najbolj umerljiv parameter 1, za najbolj negotova parametra pa označita parametra 6 in 5. Parametri so na obeh grafih razporejeni od najnižje variance do najvišje. Vidno je, da je tudi vrstni red po obeh metodah skorajda enak.

Rezultati primerjave so nad pričakovanji, saj metoda FOSM le aproksimira občutljivostno matriko. Poleg tega je dovzetna za numerične napake. Zaradi množenja variančno-kovariančne matrike z parametrom s^2 so rešitve za velikostni razred manjše od rezultatov metode na osnovi statistične analize populacije.

Za primerjavo vzemimo še primer, ko vrednosti parametrov ne zasedejo takšnih vrednosti, kot so bile navedene v tabeli 5.6. Poglejmo si nekaj primerov, kjer premaknemo enega ali več parametrov in vpliv takšnega posega na analize. V tabeli 5.7 so predstavljene statistične analize, kjer se navedenim parametrom spremeni vrednost za 1. V glavi stolpca so navedeni parametri, ki so bili pri poizkusu spremenjeni. Za primerjavo je priložen tudi stolpec analize z nespremenjenimi parametri, torej tako analizo, kot je v tabeli 5.5.



Slika 5.6: Primerjava analiz umerljivosti.

Parameter	Original	1	6	4,2
of_{flow}	0.0039	524.9086	0.1169	35.0390
$of_{pressure}$	0.0001	9.4334	0.0002	0.1186
$MSE_{flow}^{validation}$	0.0099	43.6173	0.01319	0.7126
$MSE_{pressure}^{validation}$	$3.30 * 10^{-5}$	4.7309	$4.31 * 10^{-5}$	0.2022
s^2	0.0016	203.9157	0.04402	13.2286
s_{flow}^2	$8.29 * 10^{-5}$	10.9355	0.0024	0.7299
$s_{pressure}^2$	$3.07 * 10^{-6}$	0.1965	$5.00 * 10^{-6}$	0.0024
$\hat{\Sigma}$	0.0628	12.9843	2.5657	10.2666
$BIAS_{flow}$	-0.0046	-4.6808	-0.0032	-0.1086
$BIAS_{pressure}$	0.0011	-2.7367	0.0037	-0.1306
$MAX(\Delta_{flow})$	0.1239	37.4364	0.7235	13.5800
$MIN(\Delta_{flow})$	-0.1204	-38.7429	-0.6836	-13.5888
$MAX(\Delta_{pressure})$	0.0297	-0.1141	0.0316	0.2967
$MIN(\Delta_{pressure})$	-0.0228	-5.3677	-0.0366	-1.1730
$\sigma(\Delta_{flow})$	0.0629	22.4276	0.3419	5.9183
$\sigma(\Delta_{pressure})$	0.0121	1.3940	0.0150	0.3186
$RELBIAS_{flow}$	-0.0740	-0.2087	-0.0096	-0.0183
$RELBIAS_{pressure}$	0.0963	-1.9632	0.2466	-0.40997

Tabela 5.7: Analiza umerjanj modela Anytown s spremenjenimi parametrih.

5.4 Primer: Aques de Coimbra

Primer umerjanja na modelu Aques de Coimbra je precej kompleksnejši od prej opisanega. vodovodni sistem sestavlja veliko večje število cevi in spojev. Grupiranje je bila narejena na podlagi lastnosti cevi, pri čemer ni bila upoštevana njihova povezljivost. Rezultat grupacije je 16 grup, ki so vidne na sliki 5.7.

Uporabljenih je nekoliko več merilnih mest, kot je vidno v tabeli 5.8. Delež merilnih mest je zaradi velikega števila cevi in spojev sicer precej manj ugoden od prejšnjega primera. Delež merilnih mest pretoka je 1.7%, delež merilnih mest tlaka pa 2.5%. Po drugi strani je na merilno mesto na voljo veliko več meritev, kot v primeru Anytown. Ker ta hidravlična simulacija traja 24 ur, meritve pa se izvajajo na vsakih 5 minut, je vseh časovnih korakov 288, skupno vseh meritev pa 3456. Spreminjanje pretokov merilnih mest je vidno na grafu 5.8 tlakov pa na grafu 5.9. Viden je vpliv črpalk, kot tudi nihanja vodne gladine vodohranov.

Tip merilnega mesta	Število	Merilna mesta
Meritev pretoka	5	277, 251, 303, 244, 266
Meritev tlaka	7	242, 62, 18, 69, 188, 183, 121
Verifikacijska meritev pretoka	2	295, 103
Verifikacijska meritev tlaka	6	168, 219, 142, 116, 73, 93

Tabela 5.8: Merilna mesta v modelu Aques de Coimbra.

Uporabljene nastavitve GA so identične prejšnjemu primeru zaradi primerjave delovanja. Edina sprememba je zaustavitveni pogoj, ki je v tem primeru nastavljen na $\varepsilon = 0.000001$. Algoritem se je pri tej nastavitvi zaustavil pri generaciji 450. Grafa povprečnih vrednosti ciljnih funkcij po generacijah sta značilne oblike (slika 5.10). Model se najprej hitro približa, podrobne nastavitve hrapavosti pa trajajo dlje.

Statistična analiza umerjanja v tabeli 5.9 kaže zelo visoko stopnjo ujemanja pri ciljnih funkcijah, kar je pričakovano zaradi nizko nastavljenega zaustavitvenega pogoja ε . Da se model tudi dobro ujema s situacijo v naravi, kažeta visoki ujemanji validacijskih funkcij.

Parameter	Vrednost
of_{flow}	$9.05 * 10^{-6}$
$of_{pressure}$	$1.14 * 10^{-5}$
$MSE_{flow}^{validation}$	0.0002
$MSE_{pressure}^{validation}$	$4.41 * 10^{-5}$
s^2	$1.04 * 10^{-5}$
s_{flow}^2	$2.63 * 10^{-9}$
$s_{pressure}^2$	$3.32 * 10^{-9}$
$\widehat{\Sigma}$	585372.2665
$BIAS_{flow}$	-0.0003
$BIAS_{pressure}$	0.0001
$MAX(\Delta_{flow})$	0.0066
$MIN(\Delta_{flow})$	-0.0068
$MAX(\Delta_{pressure})$	0.0091
$MIN(\Delta_{pressure})$	-0.0110
$\sigma(\Delta_{flow})$	0.0029
$\sigma(\Delta_{pressure})$	0.0033
$RELBIAS_{flow}$	-0.1165
$RELBIAS_{pressure}$	0.0538

Tabela 5.9: Analiza umerjanja modela Aques de Coimbra.

Povsem drugo sliko vidimo pri skupni negotovosti parametrov $\widehat{\Sigma}$, ki kaže na visoko nezaupanje v rezultate optimizacije. Podrobnejša analiza rezultatov v tabeli 5.10 pokaže, da k visoki vrednosti prispevata le varianci parametrov 13 in 9. Njihovi visoki vrednosti nakazujeta, da omenjenim parametrom ne gre zaupati. Po drugi strani pa so vrednosti vsekakor previsoke in so posledice deljenja z zelo majhnim številom. Omenjena numerična omejitev pri izračunu kovariančne matrike je podrobno opisana v razdelku 4.3.4. Če ti dve anomaliji zanemarimo, je splošna negotovost ostalih parametrov zelo dobra $\widehat{\Sigma} = 0.59432$, kar kaže na visoko zaupanje v ostale parametre. Analiza odstopanj kaže dobro ujemanje pri vseh meritvah, pristranskost je blizu ničle.

Grupa	Θ_i	Povprečna hrapavost	Varianca	Interval od	Interval do
12	2.4734	2.4371	$2.92 * 10^{-7}$	2.4734	2.4734
6	1.3208	1.3725	$9.11 * 10^{-6}$	1.3207	1.3208
10	1.8574	1.8256	$1.41 * 10^{-5}$	1.8573	1.8574
5	1.2667	1.1685	$2.04 * 10^{-5}$	1.2667	1.2668
11	2.1252	2.1075	$2.16 * 10^{-5}$	2.1251	2.1252
8	1.5574	1.5844	$4.12 * 10^{-5}$	1.5573	1.5575
3	1.0314	0.9398	$5.61 * 10^{-5}$	1.0313	1.0315
1	0.4481	0.01	0.0004	0.4472	0.4491
7	1.3270	1.3909	0.0006	1.3257	1.3283
4	1.1470	1.0737	0.0035	1.1401	1.1539
16	0.1943	18.5277	0.0096	0.1753	0.2133
14	4.7687	6.7474	0.1141	4.5450	4.9924
15	13.9986	16.3878	0.4656	13.0858	14.9113
9	5.4864	1.6254	1596.9459	-3124.5276	3135.5006
13	0.5840	5.739	583774.7262	-1144197.8794	1144199.0474

Tabela 5.10: Rezultati umerjanja modela Aques de Coimbra.

Pri analizi rezultatov je vidno, da so intervali zaupanja občutno preozki, saj so neposredno odvisni od parametra s^2 , ki določa varianco. Vendar lahko rečemo, da so pri nizkih variancah rezultati optimizacije precej bližje povprečni hrapavosti, kot pri tistih z visoko. To diagnostično orodje je tako pri ocenjevanju rezultatov še vedno zelo uporabno.

Na grafih 5.11 je vidna analiza umerljivosti z obema uporabljenima metodama. Obe metodi identificirata kot slabo umerljive grupe 13, 16, 9 in 14, manjše neso-glasje pa je glede grupe 15. Grafa sta urejena od leve proti desni po naraščajoči varianci. Vidno je, da metodi razvrstita parametre v nekoliko drugačnem vrstnem redu, kar je pričakovano glede na opisane omejitve obeh metod.

Po drugi strani kot senzitivne identificirata skoraj iste parametre in se ujemata pri neumerljivih. Kot diagnostični metodi, ki pomagata uporabniku pri izločevanju

parametrov z manjšim zaupanjem, sta precej uporabni, seveda pa pri velikih problemih njihovo delovanje ni popolno.

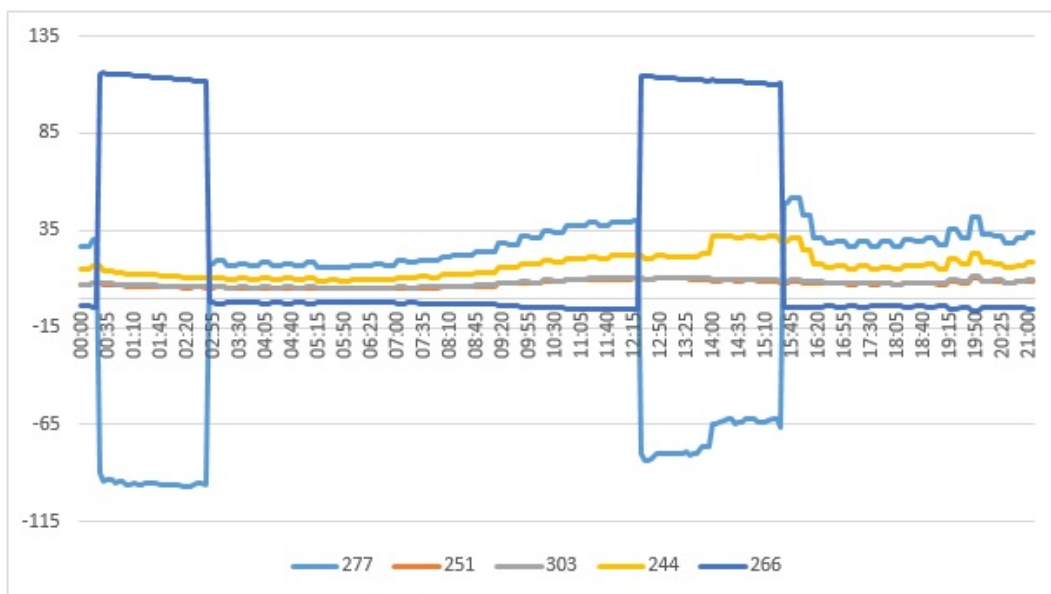
Na zadnje si pogledjmo še vplive sprememb nekaterih parametrov. Statistična analiza originalne rešitve in nekaj s spremenjenimi vrednostmi parametrov so vidne v tabeli 5.11.

Parameter	Original	5	12	15, 16
of_{flow}	$9.05 * 10^{-6}$	0.0044	75.8633	$1.10 * 10^{-5}$
$of_{pressure}$	$1.14 * 10^{-5}$	0.0010	0.1113	$1.15 * 10^{-5}$
$MSE_{flow}^{validation}$	0.0002	0.0011	94.8294	0.0002
$MSE_{pressure}^{validation}$	$4.41 * 10^{-5}$	0.0014	0.1249	$4.44 * 10^{-5}$
s^2	$1.04 * 10^{-5}$	0.0024	31.8127	$1.13 * 10^{-5}$
s_{flow}^2	$2.63 * 10^{-9}$	$1.28 * 10^{-6}$	0.0220	$3.20 * 10^{-9}$
$s_{pressure}^2$	$3.32 * 10^{-9}$	$3.06 * 10^{-7}$	$3.23 * 10^{-6}$	$3.36 * 10^{-9}$
$\widehat{\Sigma}$	585372.2665	$1.0687 * 10^7$	$3.0245 * 10^{11}$	345006.1948
$BIAS_{flow}$	-0.0003	0.0256	-0.0003	-0.0003
$BIAS_{pressure}$	0.0001	-0.0154	0.0183	-0.0001
$MAX(\Delta_{flow})$	0.0066	0.3065	119.9602	0.0110
$MIN(\Delta_{flow})$	-0.0068	-0.0090	-119.9701	-0.0113
$MAX(\Delta_{pressure})$	0.0091	0.0054	7.0941	0.0083
$MIN(\Delta_{pressure})$	-0.0110	-0.1466	-7.3521	-0.0123
$\sigma(\Delta_{flow})$	0.0029	0.0613	8.7099	0.0033
$\sigma(\Delta_{pressure})$	0.0033	0.0285	0.3331	0.0033
$RELBIAS_{flow}$	-0.1165	0.4188	$-3.90 * 10^{-5}$	-0.1096
$RELBIAS_{pressure}$	0.0538	-0.5436	0.0550	-0.03886

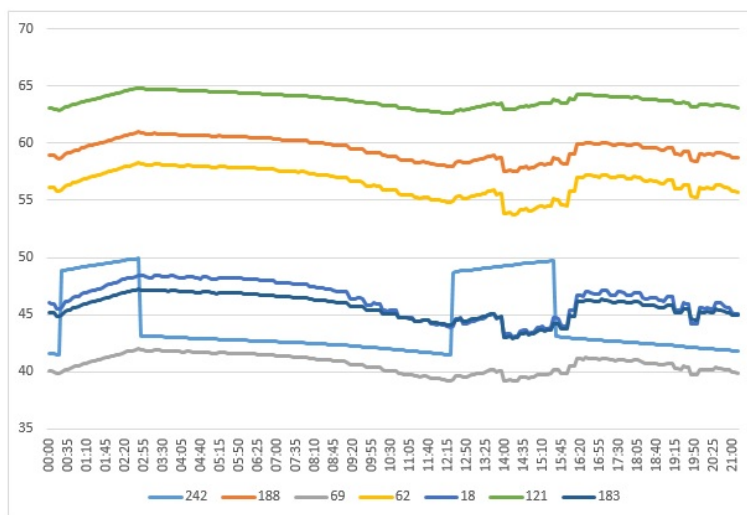
Tabela 5.11: Analiza umerjanj modela Aques de Coimbra s spremenjenimi parametrih.



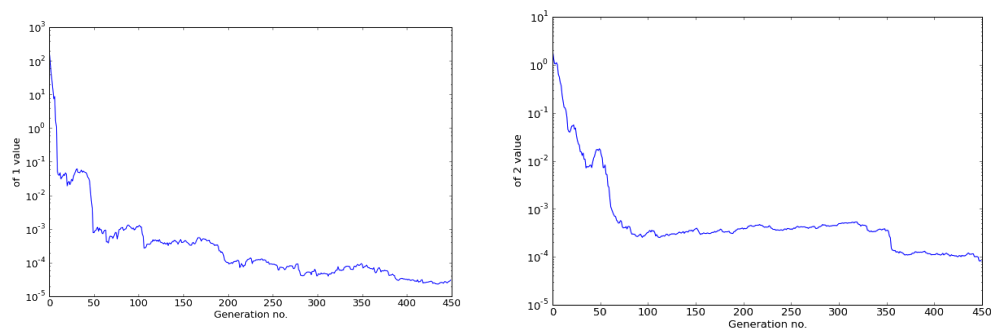
Slika 5.7: Grupacija.



Slika 5.8: Merilna mesta pretokov in njihovi podatki v različnih časovnih korakih.

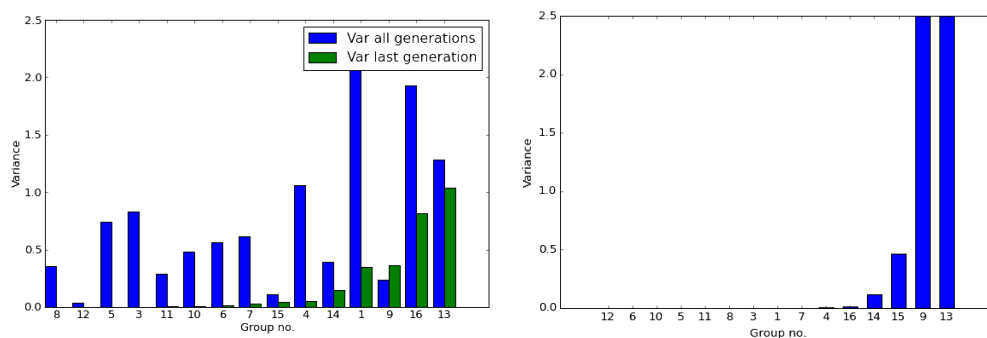


Slika 5.9: Merilna mesta tlakov in njihovi podatki v različnih časovnih korakih.



(a) Povprečna vrednost of_{flow} po generacijah. (b) Povprečna vrednost $of_{pressure}$ po generacijah.

Slika 5.10: Povprečne vrednosti ciljnih funkcij po generacijah.



(a) Variance z GA.

(b) Variance z FOSM .

Slika 5.11: Primerjava analiz umerljivosti.

5.5 Primer: Aques de Coimbra z odstranjevanjem grup

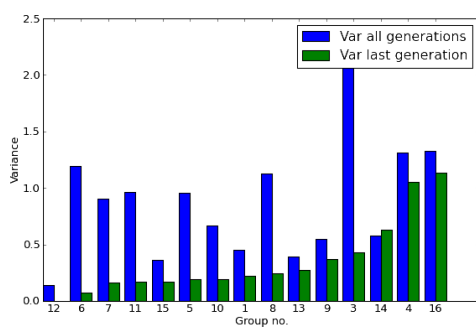
V tem razdelku testiramo način umerjanja, opisan v razdelku 4.2.3. Kot je razvidno na diagramu 4.2, optimizacijski algoritem najprej poženemo za nekaj generacij. S pomočjo analize umerljivosti in nekaj uporabnikovega vnosa odstranimo neumerljive grupe. Postopek klasično nadaljujemo do željenega zaustavitvenega kriterija. S tem se najverjetneje izboljša natančnost in verjetje v dobljene rešitve, zaradi zmanjšanja problema pa tudi hitrost konvergence.

V tem primeru bomo iskali neznane parametre že poznane modela Aques de Coimbra. Da bomo lahko hitrost in natančnost preverjali s klasičnim pristopom, je uporabljena identična grupacija kot v razdelku 5.4. Tudi merilna mesta in uporabljene meritve so identične.

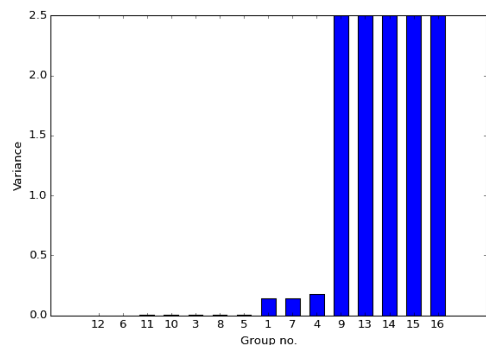
Algoritem najprej poženemo za majhno število generacij. V tem primeru smo uporabili zaustavitveni kriterij $\varepsilon = 0.1$, pri čemer se je optimizacija zaustavila pri generaciji 50. Določanje umerljivosti tako hitro, seveda ni popolno, izloči pa se vsaj večina neumerljivih parametrov. Rezultati analize umerljivosti so vidni v grafu 5.12. Vidne so razlike med omenjeno analizo in analizo po popolni konvergenci na grafu 5.11. Vseeno sta obe analizi po 50 generacijah identificirali 5 enakih slabih grup od 6 najslabših, kot analiza po popolni konvergenci. V tem trenutku odmislimo znanje iz prejšnjega primera in uporabimo samo podatke iz tega razdelka. Odrežemo najslabših 6 grup, kot so bile identificirane po 50 generacijah, torej grupe 13, 9, 1, 14, 9, 16. Algoritem poženemo naprej z enako nastavitvijo $\varepsilon = 0.000001$, kot v primeru 5.4.

Optimizacija se tokrat ustavi v generaciji 170. V primeru brez odstranjevanja grup je bilo potrebnih ob istem zaustavitvenem kriteriju 450 generacij. Statistična analiza umerjanja je vidna v tabeli 5.12. Stopnja ujemanja modela je po vseh kriterijih visoka, tudi skupna negotovost parametrov je relativno nizka, pa še to le na račun parametra 15, kot je razvidno iz grafa 5.13.

5.5. PRIMER: AQUES DE COIMBRA Z ODSTRANJEVANJEM GRUB9

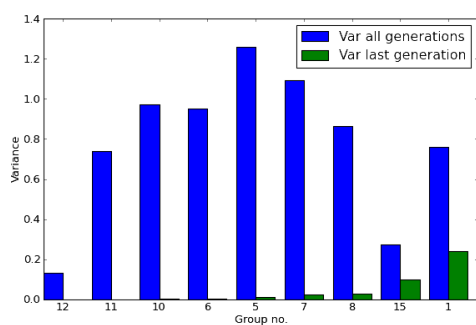


(a) Variance z GA.

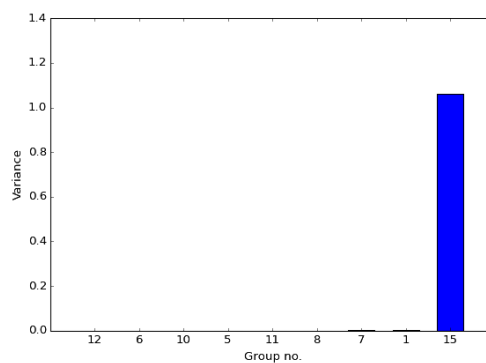


(b) Variance z FOSM .

Slika 5.12: Analiza umerljivosti pred odstranjevanjem parametrov.



(a) Variance z GA.



(b) Variance z FOSM .

Slika 5.13: Analiza umerljivosti po zaključeni kalibraciji z odstranjenimi parametri.

Parameter	Vrednost
of_{flow}	$1.11 * 10^{-5}$
$of_{pressure}$	$2.57 * 10^{-5}$
$MSE_{flow}^{validation}$	0.00026
$MSE_{pressure}^{validation}$	$4.54 * 10^{-5}$
s^2	$1.97 * 10^{-5}$
s_{flow}^2	$3.23 * 10^{-9}$
$s_{pressure}^2$	$7.47 * 10^{-9}$
$\hat{\Sigma}$	1.0647
$BIAS_{flow}$	-0.0001
$BIAS_{pressure}$	0.0001
$MAX(\Delta_{flow})$	0.0091
$MIN(\Delta_{flow})$	-0.0114
$MAX(\Delta_{pressure})$	0.0162
$MIN(\Delta_{pressure})$	-0.0141
$\sigma(\Delta_{flow})$	0.0033
$\sigma(\Delta_{pressure})$	0.0050
$RELBIAS_{flow}$	-0.0382
$RELBIAS_{pressure}$	0.0197

Tabela 5.12: Analiza umerjanja modela Aques de Coimbra.

Grupa	Θ_i	Povprečna hrapavost	Varianca	Interval od	Interval do
12	2.4556	2.4371	$5.47 * 10^{-7}$	2.4556	2.4556
6	1.4561	1.3725	$1.97 * 10^{-5}$	1.4560	1.4561
10	1.8090	1.8256	$2.56 * 10^{-5}$	1.8089	1.8090
5	1.0021	1.1685	$2.78 * 10^{-5}$	1.0020	1.0021
11	2.0560	2.1075	$3.87 * 10^{-5}$	2.0560	2.0561
8	1.5244	1.5844	$7.62 * 10^{-5}$	1.5243	1.5246
7	1.4478	1.3989	0.0014	1.4450	1.4506
1	1.9134	0.8974	0.0026	1.9082	1.9186
15	19.7787	16.3878	1.0605	17.7001	21.8573

Tabela 5.13: Rezultati umerjanja modela Aques de Coimbra z odstranjevanjem grup.

Poglavje 6

Implementacija kot GIS aplikacija

6.1 Uvod

Geografsko informacijski sistem (ang. Geographical information system) je računalniški podatkovno procesni sistem za upravljanje prostorskih (geografskih) podatkov. V strokovni javnosti se na tovrstne sisteme sklicuje kar s kratico GIS, katero bomo uporabljali tudi v nadaljevanju. Njegove glavne naloge so uvoz in shranjevanje prostorskih podatkov, njihova vizualizacija ter seveda obdelava, vzdrževanje in analiza [24].

Ker gre pri vodovodnih sistemih za podatke, katere pogosto obdelujemo znotraj GIS, so bila prej opisana orodja za umerjanje razvita znotraj enega tovrstnih sistemov.

6.2 Vrste podatkovnih tipov

V GIS so podatki združeni v t.i sloje. Sloj (ang. layer) združuje geografske objekte istega ali pa zelo podobnega tipa. Vsak objekt (ang. feature) znotraj sloja nosi seveda lokacijske podatke, ki objekt umestijo v prostor. Poleg tega mu lahko pripišemo tudi vrsto atributnih lastnosti, ki niso vezane na postavitev. Slednje so lahko najrazličnejših podatkovnih tipov; od celoštevilskih in decimalnih števil,

nizov ali daljših tekstovnih opisov.

V praksi srečamo dva osnovna tipa slojev.

1. Rasterski sloji

Rasterski sloji so podobni digitalnim fotografijam, ki jih lahko povečujemo in zmanjšujemo. Tudi v rasterskem sloju so osnovni elementi postavljeni v dvodimenzionalno mrežo. Če so v fotografiji osnovni elementi točke, so v rasterskem sloju celice. Za oboje pa velja, da so to homogene enote, ki same po sebi še ne sestavljajo oblik. Slednjo dobimo, ko združimo množico točk v mrežo.

Najbolj popularen tip rasterskih slojev so aerofotografije.

2. Vektorski sloji

Bolj pogosto si geografske objekte predstavljamo kot geometrijske oblike. Tovrstnim slojem rečemo vektorski sloji. Uporabljamo tri različne tipe tovrstnih slojev, s katerimi lahko v praksi popišemo veliko večino človeških pa tudi naravnih geografskih objektov.

(a) Točka

Točke so brezdimenzijski objekti, ki jih uporabljamo takrat, ko je objekt omejen na eno samo lokacijo. Shranjevanje takšnega objekta porabi zelo malo spomina, saj se zapiše kot par koordinat v prostoru. Z njimi ponavadi opišemo objekte, kot so prometni znaki, spoji na vodovodnem ali električnem omrežju, rezervoarji, objekti zelo majhnih dimenzij, stičišča cest (križišča) in lokacije merilnih instrumentov.

(b) Črta

Črte so enodimenzijski geografski objekti. Uporabljamo jih takrat, ko objekt lepo opišemo s črto. To so na primer ceste, železnice, mostovi, električni vodi, cevovodi, državne in občinske meje.

V podatkovnem modelu se črto predstavi kot množico (seznam) točk, ki sestavljajo lomljenko.

(c) Poligon

Poligoni so dvodimenzijski geografski objekti. Podobno kot črte jih sestavlja množica točk, ki paroma napenjajo daljice. Bistvena razlika

med črto in poligonom pa je ta, da se pri poligonu prva in zadnja daljica dotitaka in tako oblikujeta zanko. Poligon tako omejuje območje na zemljevidu z lomljeno črto.

V praksi z njim najpogosteje označujemo stanovanjske in gospodarske stavbe, igrišča, parkirišča, zaščitena območja, države in parcele. Poleg tega z njimi opišemo tudi naravne značilnosti, kot so na primer jezera ali pa naravni habitati neke ogrožene vrste.

6.3 Vzpostavitev hidravličnega modela

Hidravlični model vzpostavimo na podlagi podatkov iz vektorskih slojev in nekaj atributnih tabel. Njihova struktura natančno sovпада s strukturo elementov vodovodnega omrežja, kot so opisani v razdelku 2.2. Za vzpostavitev osnovnega hidravličnega modela je bila v tej nalogi med drugim uporabljen program Oslandia. Oslandia [23] je GIS-ovski vtičnik, ki omenjene podatke prebere in v ustrezni obliki zapiše v tekstovno .inp datoteko, ki jo lahko v nadaljevanju bere in uporablja EPANET2.

6.4 QGIS

QGIS je odprtokodno GIS-ovsko okolje, znotraj katerega so bile implementirane funkcionalnosti, opisane v prejšnjih poglavjih, za umerjanje vodovodnih sistemov. Okolje je dostopno pod GNU General public licence [25]. Razvijajo ga prostovoljci, skupnost tako skrbi za razvoj in odpravo napak ter pomoč uporabnikom. Okolje se neprestano širi, saj lahko uporabniki po želji okolje izpopolnjujejo. Enostaven način razširitve je z dodatnimi programi - vtičniki. Slednji dostopajo in manipulirajo z geografskimi podatki v prej omenjenih slojih. Na ta način lahko razvijamo specializirana orodja, hkrati pa ohranimo splošne funkcionalnosti, kot so urejanje slojev in vizualizacija geografskih podatkov.

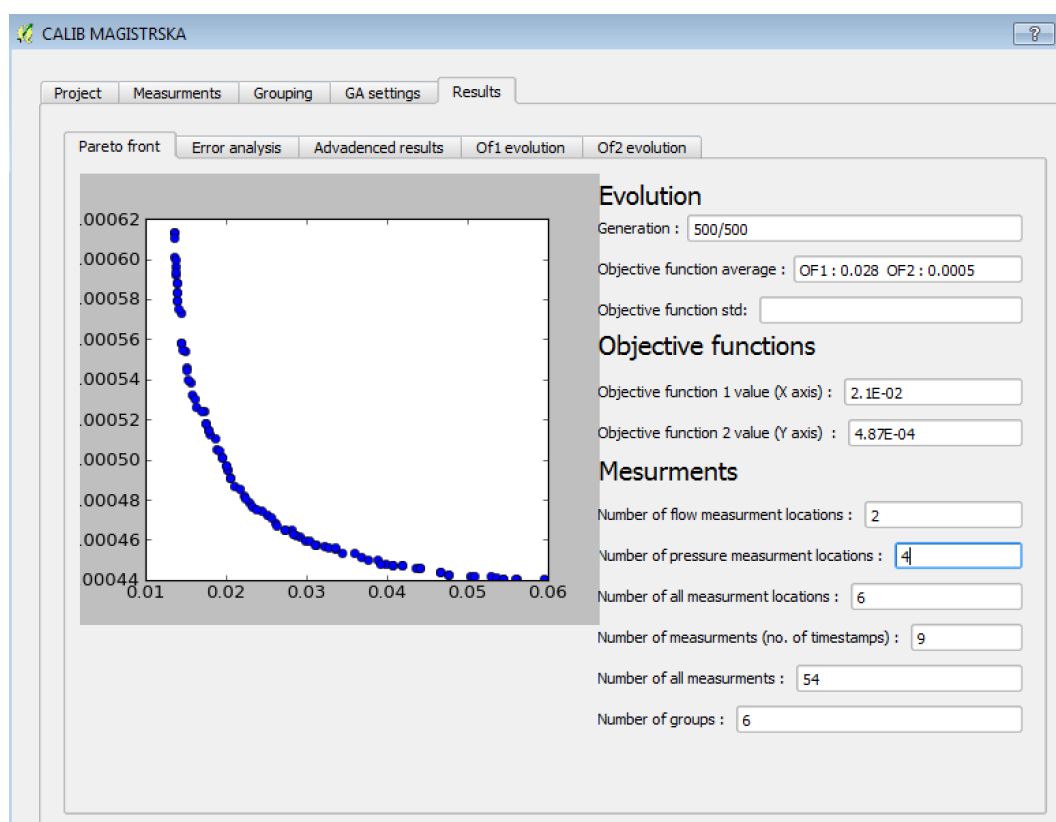
6.5 Vtičnik CALIB

Vtičnik CALIB je orodje, implementirano znotraj prej omenjenega okolja QGIS in je razvito kot povezovalni del te naloge. Znotraj vtičnika so implementirane vse opisane funkcionalnosti, pa tudi nekaj takih, ki so potrebne za delovanje, a jih nismo posebej omenili.

Vtičnik je napisan v programskem jeziku Python, uporabljene so dodatne knjižnice, kot na primer EPANET za delo s hidravličnimi modeli in Matplotlib za izris grafov. Uporabniški grafični vmesnik je oblikovan z orodjem QT.

Orodje najlažje opišemo po sklopih oz. zavihkih, ki jih predstavljajo v uporabniškem grafičnem vmesniku.

1. **Projekt** V tem zavihku uporabnik izbere obstoječi ali ustvari nov kalibracijski projekt, izbere ustrezne sloje s podatki in uvozi .imp datoteko.
2. **Meritve** V zavihku z meritvami se uvozijo podatki tlakov in pretokov za umerjanje ter validacijo v .csv obliki.
3. **Grupe** V tem delu imamo pregled nad trenutno razvrstitvijo cevi v grupe. Po želji se lahko grupacija izvede ponovno glede na različne kriterije. Po končani optimizaciji se pokaže tudi grafi varianc, uporabnik pa lahko za željeno grupo izriše histogram, tako zadnje kot vseh generiranih populacij.
4. **Genetski algoritem** V tem zavihku se nastavijo parametri optimizacije in prikažejo statusna sporočila. Ozadje tega dela programa je obsežno, saj vključuje celoten optimizacijski algoritem.
5. **Rezultati** Ta del programa je najobsežnejši predvsem vizualno. Prikazuje celotno populacijo tako med optimizacijo, kjer lahko opazujemo premikanje osebkov, kot tudi po zaključku delovanja GA. Ta del aplikacije je viden na sliki 6.1. Uporabnik s klikom na željeni osebek prikaže podatke izbrane rešitve ter vse statistične analize, ki smo jih uporabili v nalogi. V tem zavihku tudi prikažemo ostale pomožne grafe, kot so npr. grafi povprečnih vrednosti ciljnih funkcij po generacijah.



Slika 6.1: Vtičnik CALIB postavljen v zavihek z rezultati.

Poglavje 7

Zaključek

Naloga bralcu iz računalniške stroke, brez predhodnega znanja o mehaniki tekočin predstavi problem umerjanja parametrov vodooskrbnih sistemov. Predvsem uvodno poglavje o hidravličnih modelih 2 je namenjeno nepoučenemu bralcu in je kot tako dobro izhodišče za prihodnje delo na interdisciplinarnem področju mehanike tekočin in računalništva. V zaključnem delu tega poglavja je bila tudi izpeljana časovna analiza hidravlične simulacije s stalnim tokom, katere ni še na voljo v literaturi.

V poglavju o genetskih algoritmih 3 se lahko bralec pouči o omenjenem pristopu k reševanju časovno zelo zahtevnih problemov. Podrobno je bil v ta namen opisan algoritem NSGA-II, ter opravljene izpeljave časovnih analiz. Za piko na i je v zaključnem delu tega poglavja opisanih tudi več pristopov za manipulacijo z populacijo in osebki ter predstavljene dobre in slabe strani navedenih pristopov.

Praktični del naloge se začne v poglavju o umerjanju parametrov vodovodnih sistemov 4. Za namen te naloge je bil razvit osrednji algoritem, ki s pomočjo večkriterijske optimizacije išče vrednosti parametrov vodovodnega sistema, ter tako model čim bolj približa situaciji v naravi. Velik poudarek je bil namenjen tudi diagnostičnim orodjem, namenjenim analizi umerjanja, kjer se sprašujemo o kvaliteti rezultatov algoritma. Metoda analize umerljivosti na podlagi statistične analize populacije je eno takih orodij in je bila v tej nalogi predstavljena kot alternativna rešitev uveljavljeni metodi FOSM. Izkazalo se je, da imata obe metodi analize umerljivosti določene pomankljivosti, vseeno pa sta dobri diagnostični orodji in tako praktično uporabni. V tem poglavju je sicer opisanih še precej

drugih analitičnih orodij, ki so bila razvita v sklopu te naloge.

Program, opisan v poglavju 4, smo testirali na več praktičnih primerih v predzadnjem poglavju 5. Kot osnova sta bila uporabljena dva različna hidravlična modela. Izkazalo se je, da je metoda umerjanja parametrov vodovodnih sistemov s pomočjo genetskih algoritmov dobro orodje. V vseh primerih smo v relativno majhnem številu generacij dobili vrednosti parametrov, ki se dobro ujemajo z realno situacijo, kar so pokazale prav vse analitične metode. Največje nesoglasje je bilo po pričakovanjih pri analizi umerljivosti parametrov. Navedene teoretične omejitve obeh metod, opisanih v poglavju 4, so se izkazale za resnične, vseeno pa sta obe dobro diagnosticirali problematične, neumerljive parametre.

Zadnje poglavje 6 je bolj inženirsko. Vsa orodja, opisana znotraj te naloge so bila implementirana znotraj geografsko informacijskega sistema, kot razširitev okolja QGIS v obliki vtičnika. Na ta način je umerjanje parametrov uporabniku bolj prijazno. Tako uporabnik z nekaj kliki v aplikaciji naloži podatke meritev in uvozi hidravlični model. Tudi ostale funkcionalnosti, kot so nastavitve genetskega algoritma in analiza rezultatov, so vključene v omenjeni vtičnik. Grafi, vidni v tej nalogi, so tako zaslonske slike aplikacije. Ker gre pri implementaciji za strokovno in ne znanstveno področje, se temu delu naloge ni posvečalo velike pozornosti.

7.1 Nadaljnje delo

Umerjanje parametrov vodovodnih sistemov je široko področje, v katerega so raziskovalci skozi leta vključili precej truda, poglavje pa s tem še ni končano.

Izboljšave so vedno možne, zato bomo na tem mestu navedli nekaj možnih nadaljevanj te naloge. Kot enotavnejše izboljšave dela v tej nalogi bi lahko navedli še kup drugih diagnostičnih orodij, ki niso bila implementirana. GIS-ovsko okolje po drugi strani omogoča vrsto dodatnih vizualizacij umerjanja, kot na primer prikaz negotovosti na sloju s cevmi ali pa vrednosti hrapavosti.

Bolj dolgoročne razširitve dela v tej nalogi so povezane s področjem IoT. V nalogi je bilo pokazano, da lahko razvijemo hidravlični model, ki z veliko natančnostjo simulira dogajanje v vodovodnem sistemu. Meritve, vzete v realnem času, bi lahko tako uporabili in opravljali hkratno realnočasovno kalibracijo in diagnostiko sistema. Takšen program bi na primer prekomerno porabo vode zaznal

kot odstopanje in upravljalcu sistema javil nepričakovano dogajanje. Na tak način bi lahko napake, kot je puščanje cevovoda, lahko zaznali zelo hitro.

Literatura

- [1] Water supply and distribution modeling DHL technology flyer Dostopno na: <http://www.dhigroup.com/upload/publications/scribd/166953527-Water-Supply-and-Distribution-Modelling-DHI-Technology-Flyer.pdf> (pridobljeno 22.7.2016)
- [2] D. A. Savic and Z. S. Kapelan and P. M. Jonkergouw, "Quo vadis water distribution model calibration?", Urban Water Journal, št. 6, str. 3 - 22, 2009
- [3] C. J. Hutton, Z. Kapelan, L. Vamvakieridou-Lyroudia in D. A. Savic , "Dealing with Uncertainty in Water Distribution System Models: A Framework for Real-Time Modeling and Data Assimilation", JOURNAL OF WATER RESOURCES PLANNING AND MANAGEMENT, str. 169 - 183, 2014
- [4] N. Fontanaa and M. Giugnib and S. Gliozzic and M. Vitalettic, "Shortest path criterion for sampling design of water distribution networks", Urban Water Journal, št. 12, str. 154 - 164, 2015
- [5] Epanet users manual Dostopno na: <http://www.innovyze.com/products/epanet/download/P1007WWU.pdf> (pridobljeno 22.7.2016)
- [6] EPANET Program Methodology Dostopno na: http://www.hydroasia.org/jahia/webdav/site/hydroasia/shared/COURSES/MANUALS/DHI_Urban_Software/MIKE_NET-Water_Distribution_Networks/MIKE_NET_Methodology.pdf (pridobljeno 25.7.2016)
- [7] D. Kozelj, "Večkriterijska optimizacija umerjanja hidravličnih modelov vodovodnih sistemov", Doktorska disertacija, Fakulteta za gradbeništvo in geodezijo, Univerza v Ljubljani, 2014

-
- [8] Wikipedija: Kirchhoff's circuit laws Dostopno na: https://en.wikipedia.org/wiki/Kirchhoff%27s_circuit_laws (pridobljeno 8.8.2016)
- [9] Paul F. Boulos, Kevin E. Lansey, Bryan W. Karney, "Comprehensive Water Distribution Systems Analysis Handbook for Engineers and Planners", American Water Works Assn, 2006, pogl. 5
- [10] Iftekhhar Ahmed, "Application of the Gradient Method for the Analysis of Unsteady Flow in Water Networks", Diplomaska naloga, Faculty of civil engineering, University of Arizona, 1997
- [11] Wikipedija: Evolutionary algorithm Dostopno na: https://en.wikipedia.org/wiki/Evolutionary_algorithm (pridobljeno 5.7.2016)
- [12] Wikipedija: Pareto efficiency Dostopno na: https://en.wikipedia.org/wiki/Pareto_efficiency (pridobljeno 26.11.2016)
- [13] Melanie Mitchell, "An Introduction to Genetic Algorithms", The MIT Press, 1998
- [14] Kalyanmoy Deb, Samir Agrawal, Amrit Pratap, and T Meyarivan, "A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II", IEEE Transactions on Evolutionary Computation, št. 6, zv. 2, str. 182 - 197, 2002
- [15] Brad L. Miller, David E. Goldberg "Genetic Algorithms, Tournament Selection and the Effects of Noise", Complex Systems, št. 9, str. 193- 212, 1995.
- [16] Noraini Mohd Razali, John Geraghty, "Genetic Algorithm Performance with Different Selection Strategies in Solving TSP", v zborniku: Proceedings of the World Congress on Engineering 2011, London, 2011
- [17] Fundamentals of Genetic algorithms: AI course Lecture 39 - 40 Dostopno na: http://www.myreaders.info/09_Genetic_Algorithms.pdf (pridobljeno 25.11.2016)
- [18] Wikipedija: Newton's method Dostopno na: https://en.wikipedia.org/wiki/Newton%27s_method#k_variables.2C_m_equations.2C_with_m_.3E_k (pridobljeno 12.8.2016)

-
- [19] Wikipedija: Qualitative reasoning Dostopno na: https://en.wikipedia.org/wiki/Qualitative_reasoning (pridobljeno 28.11.2016)
- [20] Wikipedija: Variance Dostopno na: <https://en.wikipedia.org/wiki/Variance> (pridobljeno 28.10.2016)
- [21] Minitab: What is the variance-covariance matrix Dostopno na: <http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/anova/anova-statistics/what-is-the-variance-covariance-matrix/> (pridobljeno dne 28.10.2016)
- [22] Oslandia cookbook Dostopno na: https://github.com/Oslandia/qgis-water-management-cookbook/blob/master/aguas_de_coimbra/README.md (pridobljeno 4.11.2016)
- [23] Oslandia: Github Dostopno na: <https://github.com/Oslandia/qgis-epanet> (pridobljeno 8.11.2016)
- [24] Wikipedija: Geographic information system Dostopno na: https://en.wikipedia.org/wiki/Geographic_information_system, (pridobljeno 10.11.2016)
- [25] Discover QGIS Dostopno na: <http://www.qgis.org/en/site/about/index.html> (pridobljeno 10.11.2016)